



Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Estatística - EST

Análise da Série Temporal de Id 1936

Aluno Responsável:

Mathews de Noronha Silveira Lisboa - 22/0006172

Professor:

José Augusto Fiorucci

23 de janeiro de 2023

Sumário

1	Introdução	3
1.1	Transformações de Box-Cox	4
2	Seleção do Modelo	5
2.1	Seleção do Modelos Arima	5
2.2	Análise de Resíduos Arima	6
2.3	Seleção de Modelos ETS	9
2.4	Análise de Resíduos ETS	11
3	Resultados	13
3.1	Desempenho preditivo por janelas deslizantes	13
3.2	Previsões pontuais e intervalares	14
3.2.1	Benchmarks	16
4	Conclusão	18
A	Tabela de Previsões Pontuais e Intervalares	19
B	Códigos em R	20

1 Introdução

Começamos com a apresentação da série temporal escolhida dentro o banco de dados M3. A série escolhida é a de id 1936, é uma série de período mensal que compreende de janeiro de 1983 e julho de 1992, contendo 115 observações. Pode-se observar a série na Figura 1:

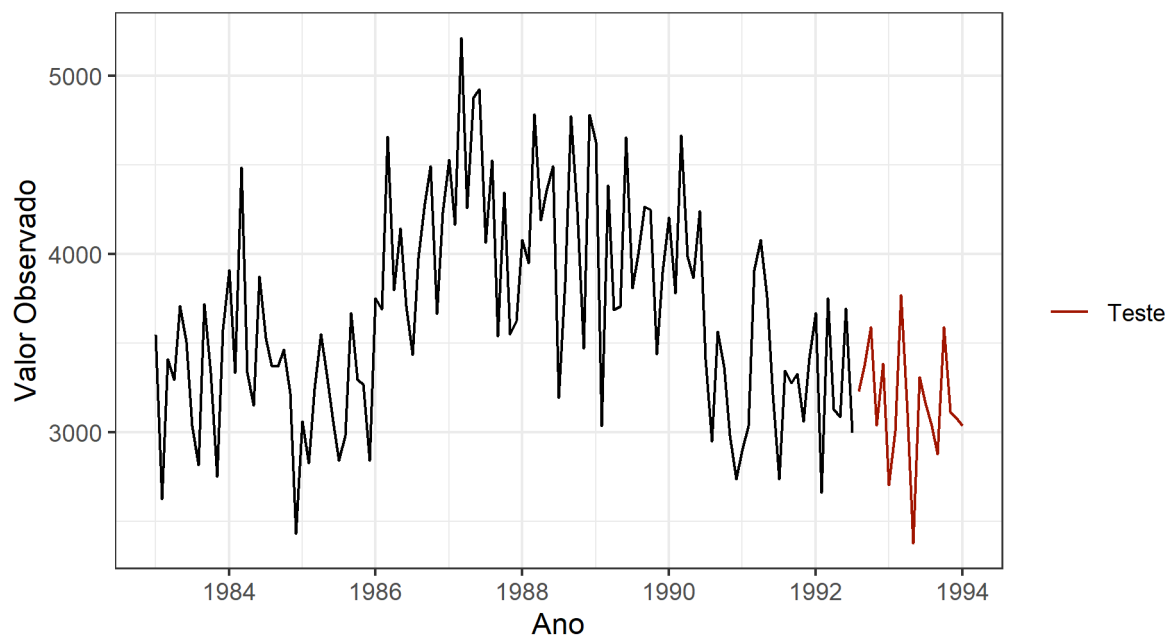


Figura 1: Apresentação da Série Original

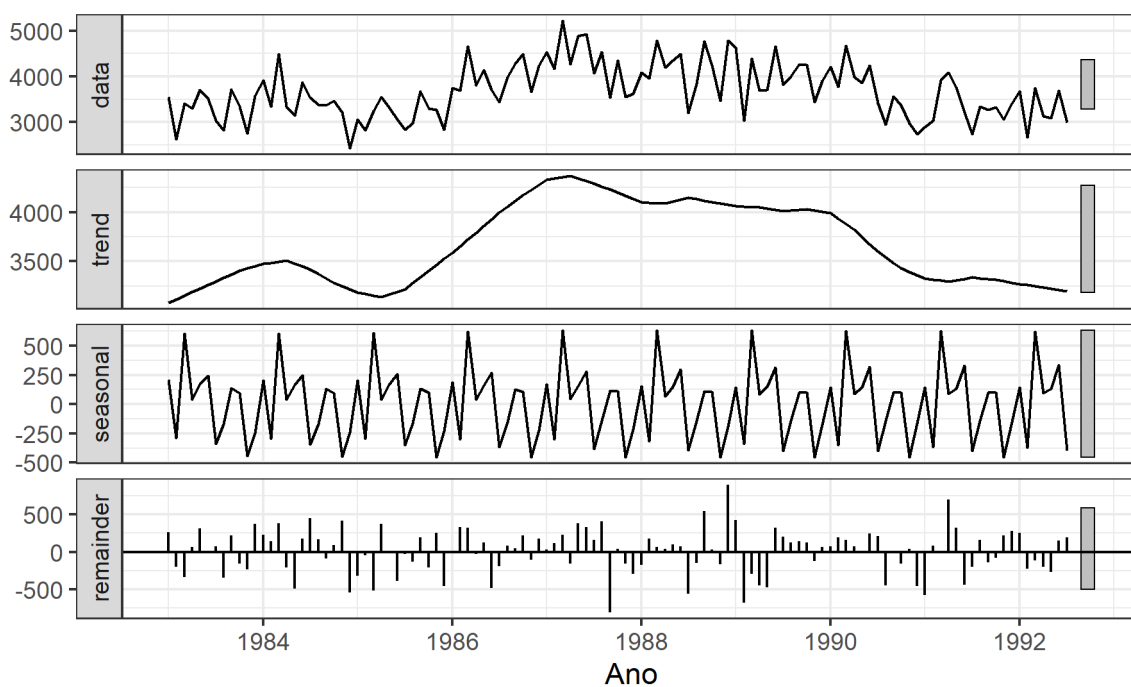


Figura 2: Decomposição da Série id 1936

Observa-se na Figura 1 que há um componente forte de tendência claro na série, porém não fica claro a respeito da sazonalidade e é difícil identificar o formato da tendência ou da série sem esse aspecto, portanto foi feito a decomposição STL da série e assim pode-se identificar mais elementos.

Então na Figura 2 mostra evidente a tendência da série, como por exemplo o crescimento entre 1984 e 1987 seguido pela queda em 1990 e 1992. Também fica evidente a sazonalidade a cada 12 meses. Deve-se dizer que claramente a série precisa de pelo menos uma diferenciação. Todas as análises e gráficos foram gerados por meio do software R.

1.1 Transformações de Box-Cox

Considere que os dados apresentam diferentes padrões de variação ao longo do tempo, sendo assim pode ser útil uma transformação nos dados. Entre transformações mais comuns estão funções positivas e inversíveis como a raiz quadrada e logaritmo. Uma gama de transformações amplamente utilizada é conhecida como transformação de Box-Cox e ela se dá por meio de:

$$w_t^\lambda = \begin{cases} \log(y_t), & \text{se } \lambda = 0; \\ \frac{y_t^\lambda - 1}{\lambda}, & \text{se } \lambda \neq 0. \end{cases} \quad (1.1.1)$$

Sendo assim, observe que se $\lambda = 1$ então é equivalente a trabalhar com a série original com uma constante que decresce em 1 em toda a série. Uma vez que for calculada uma transformação e utilizada para treinar e obter as previsões, é importante calcular a transformação inversa das previsões para retornar a escala original dos dados. O inverso da transformação de Box-Cox é obtido simplesmente por :

$$y_t = \begin{cases} \exp(w_t), & \text{se } \lambda = 0; \\ (\lambda w_t + 1)^{1/\lambda}, & \text{se } \lambda \neq 0. \end{cases} \quad (1.1.2)$$

2 Seleção do Modelo

Nessa seção vai ser feito a seleção e análise de resíduos dos modelos arima e ets. Como na apresentação da série foi encontrado um processo que possui tendência e sazonalidade faz sentido supor um modelo $SARIMA(p, d, q)x(P, D, Q)_{12}$, sendo assim é preciso encontrar os parâmetros p, d, q, P, D e Q que melhor ajustariam o processo escolhido.

2.1 Seleção do Modelos Arima

Começando pela seleção do modelo, primeiramente devemos identificar quantas diferenças precisam ser feitas, uma vez que é claro que a série original não será estacionária, uma vez que possui tendência e sazonalidade muito claras. Para isso foi feito o teste de KPSS.

Para o caso da série original o teste obtém um p.valor de 0.061 aproximadamente, rejeitando a hipótese nula no para um nível de significância de 10%. Então tomando uma diferença simples foi realizado o teste novamente e agora o resultado foi de 0.1, lembrando que o teste retorna com p.valor igual a 0.1 mesmo quando é maior que o valor impresso, dessa forma parece não ter razão para rejeitar a hipótese nula. Também foi feito o uso da função `ndiff()` e `nsdiff()` do pacote `forecast` no software R, depois da primeira diferença o resultado encontrado foi que não devia-se tirar mais nenhuma diferença simples ou sazonal. Então, sendo assim é interessante notar que já foi definido pelo menos dois parâmetros do modelo no caso os de diferença: $d = 1$ e $D = 0$.

O próximo passo lógico é realizar os gráficos de ACF e PACF, para tentar identificar alguns valores para os demais parâmetros.

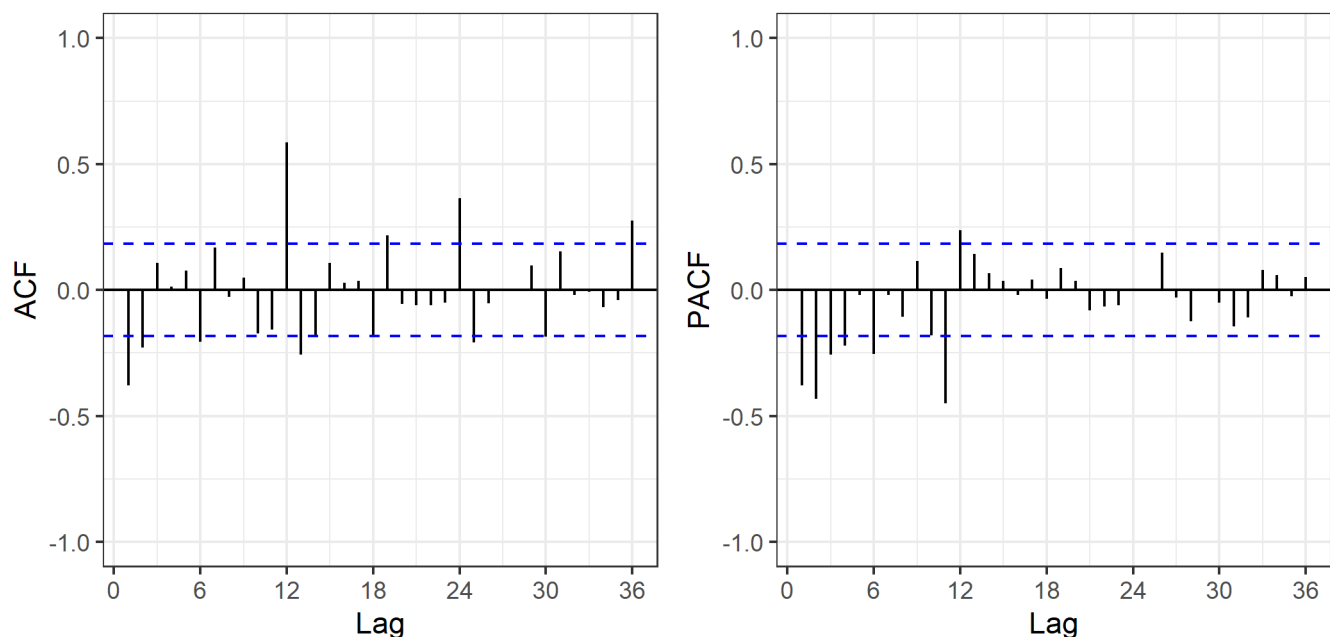


Figura 3: Gráficos ACF e PACF para a série id 1936

Aqui é interessante mostrar que observar que na Figura 3 no caso de ACF temos que tem valores representativos para lags que correspondem aos períodos de sazonalidade, observe 12, 24 e 36. Também é importante mostrar que no caso do PACF existe um corte para em torno de zero depois do segundo lag, mesmo tendo um valor representativo é apenas um depois do segundo. Sendo assim, pode-se imaginar valores como de 0 a 2 para p e q , e valores de 1 a 2 para P e Q . Assim testando algumas modelos candidatos na seleção manual tem-se o seguinte resultado:

Tabela 1: Modelos SARIMA Candidatos

Nome do Modelo	AICc	Modelo
fit_1	1709.94	$SARIMA(0, 1, 1) \times (1, 0, 2)_{12}$
fit_2	1710.05	$SARIMA(0, 1, 1) \times (1, 0, 1)_{12}$
fit_3	1745.21	$SARIMA(0, 1, 0) \times (1, 0, 1)_{12}$
fit_4	1712.11	$SARIMA(0, 1, 2) \times (1, 0, 2)_{12}$
fit_5	1711.72	$SARIMA(0, 1, 2) \times (1, 0, 1)_{12}$

Considerando a métrica do AIC corrigido denominado como AICc, escolhe o modelo que tiver o menor valor, nesse caso o modelo escolhido seria o modelo fit_1 , de configuração $SARIMA(0, 1, 1) \times (1, 0, 2)_{12}$. Também pode-se fazer a escolha do modelo de forma automática utilizando da função do R `forecast auto.arima`, nesse caso o modelo encontrado pela função automática do R foi $SARIMA(0, 1, 1) \times (0, 0, 2)_{12}$ e obteve o AICc de 1731.8, considerando essa métrica, pode-se dizer que o modelo selecionado manualmente é preferível ao que foi selecionado automaticamente.

Outra investigação possível é a respeito do modelo treinado utilizando a transformação Box-Cox. Utilizando da função para encontrar um λ adequado encontra-se $\lambda = 0.222$ e aplicando a transformação temos uma nova série w_t . O modelo selecionado para o caso da transformação de Box-Cox foi o mesmo do que foi selecionado pelo método automático $SARIMA(0, 1, 1) \times (0, 0, 2)_{12}$ com AICc de 277.38, porém observe que agora é uma série de escala menor que a original.

2.2 Análise de Resíduos Arima

Análise de resíduos é um passo crucial para que certifiquemos que os resultados encontrados são válidos, pois cada modelo tem seus pressupostos que devem ser atendidos para validar as descobertas. Dito isso, existem formas gráficas de verificar os resíduos dos modelos selecionados. Outra maneira é realizando testes para verificar estacionaridade, independência e normalidade dos resíduos, esse último é importante caso seja utilizado da distribuição normal para os intervalos de confiança.

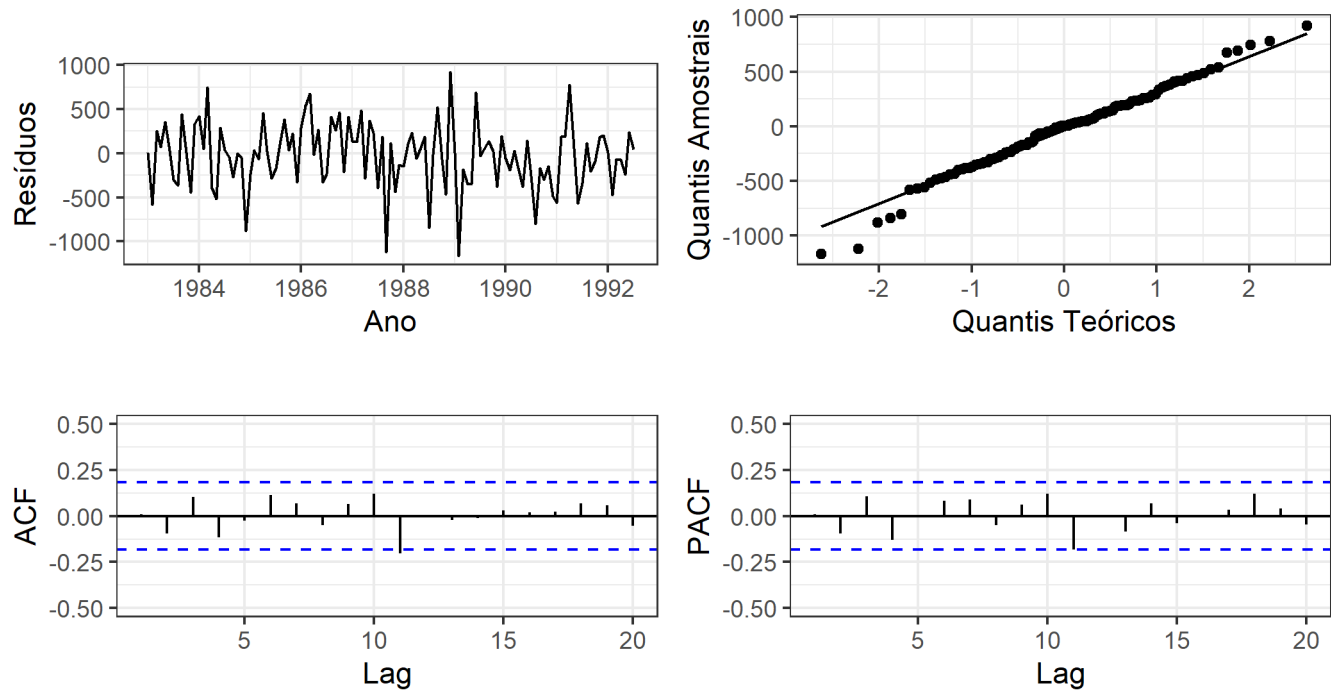


Figura 4: Gráficos de resíduos modelo $SARIMA(0, 1, 1) \times (1, 0, 2)_{12}$

Tabela 2: Resultado dos teste para modelo $SARIMA(0, 1, 1) \times (1, 0, 2)_{12}$

Teste	P.Valor
<i>KPSS</i>	0,10
<i>Ljung – Box</i>	0,70
<i>Shapiro – Wilk</i>	0,24

A Figura 4 possui 4 imagens para análise de resíduos, primeiramente observamos a série de resíduos que deve ter as características de uma séries estacionária. E de fato, pode-se observar no caso a série tem valores em torno de zero, uma variância homogênea e sem tendência. Outro gráfico na imagem é o de Q-Qplot para os quantis teóricos da normal e apesar de alguns postos não se encontrarem na reta, parece que razoável o ajuste da normal para os resíduos. E por fim dois gráficos na parte inferior da figura, no qual tem-se ACF e PACF para os resíduos com o objetivo de identificar autocorrelação, no caso não parece ter nenhuma autocorrelação relevante para o caso dos resíduos. Além da análise gráfica foi feito também testes, começando com o de estacionaridade de o kpss teve p.valor de 0.1, ou seja não há razão para rejeitar a hipótese nula. Em seguida foi feito o teste para independência, utilizando o teste de Ljung-box obteve o valor de 0.7 no p.valor indicando novamente que não deve rejeitar a hipótese nula. E por fim o teste de Shapiro-Wilk para normalidade o qual teve o p.valor de 0.24, dessa forma não rejeita-se a hipótese nula de que os dados seguem uma distribuição normal.

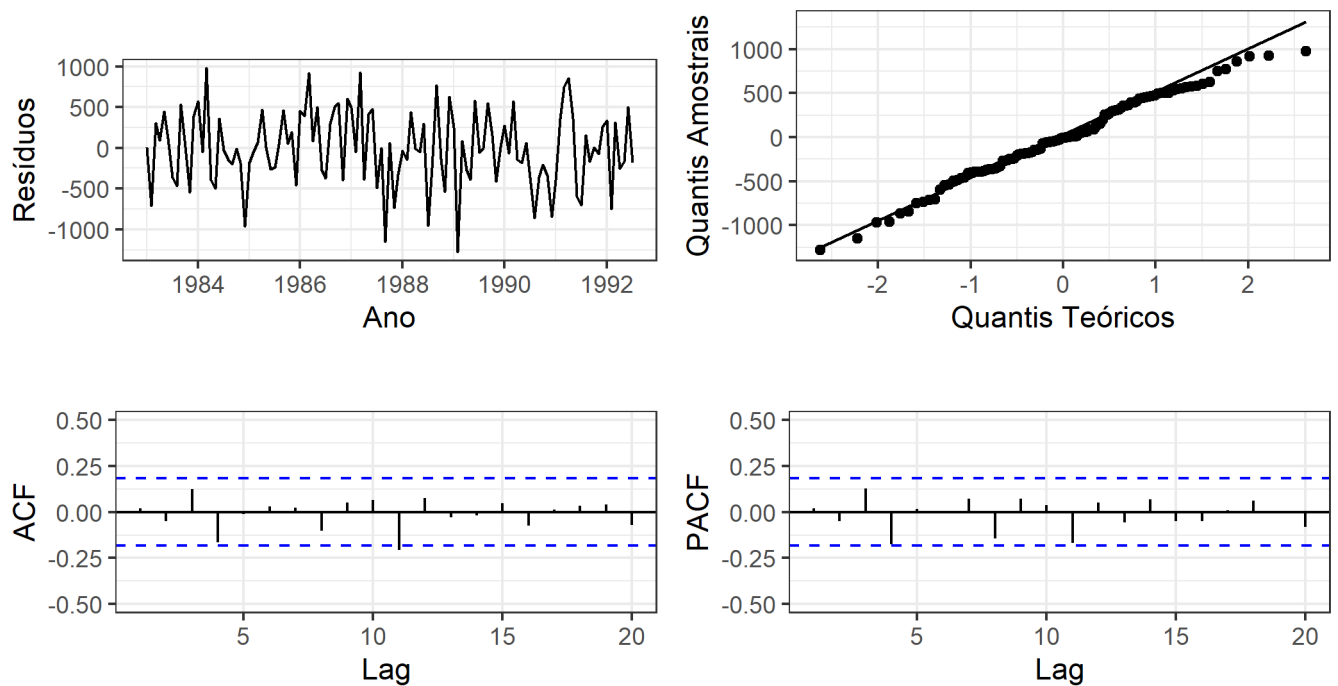


Figura 5: Gráficos de resíduos modelo automático $SARIMA(0, 1, 1) \times (0, 0, 2)_{12}$

Tabela 3: Resultado dos teste para modelo automático $SARIMA(0, 1, 1) \times (0, 0, 2)_{12}$

Teste	P.Valor
<i>KPSS</i>	0,10
<i>Ljung – Box</i>	0,65
<i>Shapiro – Wilk</i>	0,45

Ao observar a Figura 5 apresenta-se os mesmos 4 gráficos que anteriormente e novamente parece razoável a estacionaridade dos resíduos uma vez que o primeiro gráfico superior apresenta uma série em torno de zero e com variância homogênea. A respeito do gráfico de quantis teóricos apresenta-se também boa aderência a reta, indicando normalidade dos dados. E por fim quando se trata da autocorrelação apresenta-se nenhuma autocorrelação significativa para nenhum lag. Novamente foi feito o teste KPSS para estacionaridade que resultou com um p.valor de 0.1, ou seja não há razão para rejeitar a hipótese nula. Também foi feito o teste de Ljung-Box que resultou em um p.valor calculado de 0.65, logo não há razões para rejeitar hipótese nula. Também foi feito o teste de shapiro-wilk para normalidade, nesse caso o p.valor foi de 0.45, ou seja não há razão para rejeitar a hipótese nula de normalidade dos resíduos.

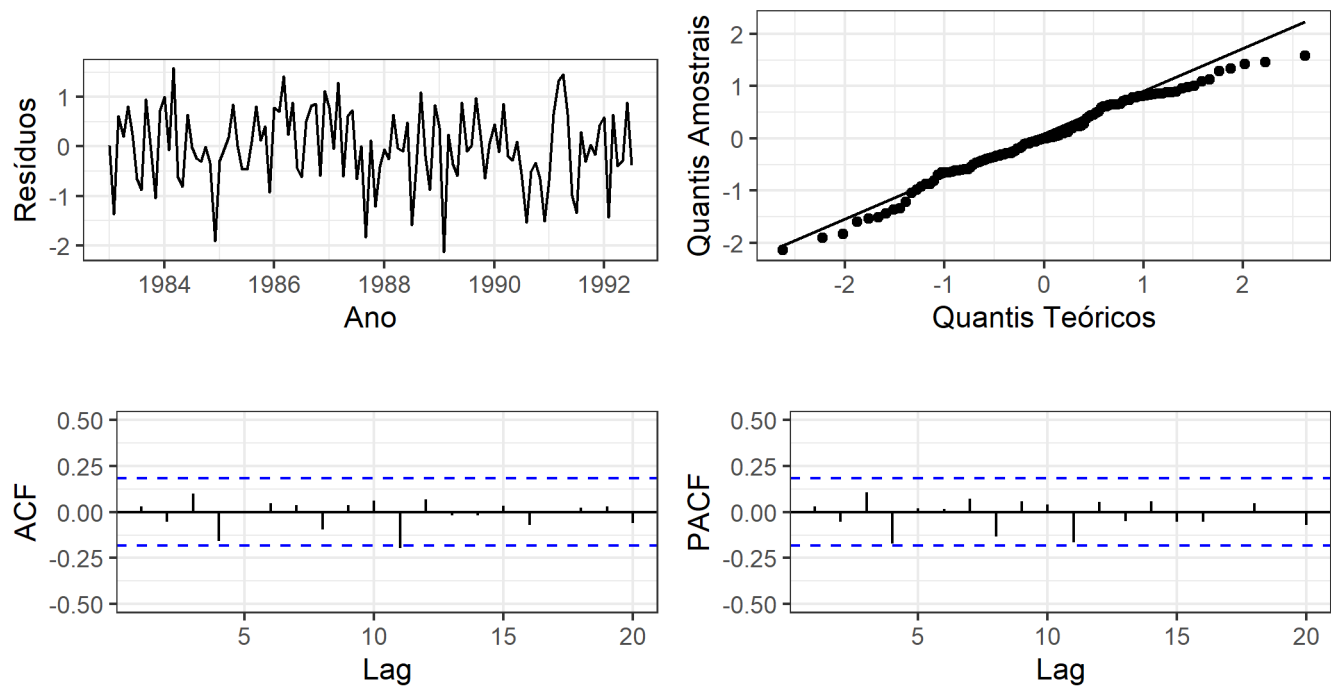


Figura 6: Gráficos de resíduos modelo transformação de Box-Cox $SARIMA(0, 1, 1) \times (1, 0, 2)_{12}$

Tabela 4: Resultado dos teste para transformação de Box-Cox $SARIMA(0, 1, 1) \times (1, 0, 2)_{12}$

Teste	P.Valor
<i>KPSS</i>	0,10
<i>Ljung – Box</i>	0,81
<i>Shapiro – Wilk</i>	0,12

Ao observar a Figura 6 mais uma vez encontra-se uma série que apresenta características de estacionaridade, com dados em torno de zero e variância homogeneia e isso se confirma quando observa-se os resultados do teste de KPSS com resultado de 0,10, ou seja é razoável afirmar é estacionário. No canto superior direito da Figura 6 observa-se o gráfico com os quantis teóricos para normalidade, diferente das outras ocasiões a aderência nas caldas não parecem estar muito boa, porém ainda é um ajuste razoável. Ao se fazer o teste de Shapiro-Wilk observa-se um p.valor de 0.12, que a nível de significância de 10%

2.3 Seleção de Modelos ETS

Quando observado na introdução que o processo apresenta tendência e também sazonalidade, logo os modelos ETS que devem ser considerados deve levar em consideração a presença dessas características.

Tabela 5: Seleção de Modelos ETS

aics	aiccs	bics	modelos	mods
1929.77	1936.89	1979.18	MAdM	10
1929.77	1936.90	1979.18	MMdM	12
1932.72	1939.03	1979.38	AAA	1
1932.82	1939.13	1979.48	MMA	7
1932.84	1939.15	1979.50	AMA	3
1933.21	1939.52	1979.88	MAA	5
1934.01	1940.32	1980.67	MMM	11
1933.67	1940.80	1983.08	MAdA	6
1933.83	1940.95	1983.24	AAAdA	2
1933.91	1941.03	1983.32	MMdA	8
1934.04	1941.16	1983.45	AMdA	4
1937.97	1944.28	1984.64	MAM	9

Observando a Tabela 5 e considerando o critério do AICc os modelos que devem ser selecionados são os ETS(MAdM) e ETS(MMdM) ou modelos 10 e 12 respectivamente. Nesse caso fica claro que ambos os modelos tem o componente de erro multiplicativo, assim como o componente de sazonalidade também é multiplicativo. A diferença entre os dois modelos está justamente no componente de tendência enquanto no modelo 10 apresenta-se com o componente de tendência aditivo amortecido enquanto no modelo 12 apresenta-se o componente de tendência multiplicativo amortecido.

Levando em consideração as mesmas observações do que foi feito para a seleção dos modelos ETS da série original, foi feito a mesma metodologia porém utilizando transformação Box-Cox com lambda igual a 0.22 .

Tabela 6: Seleção de Modelos ETS transformação de Box-Cox

aics	aiccs	bics	modelos	mods
466.46	472.76	513.12	AMA	3
467.21	473.52	513.87	AAA	1
466.46	473.58	515.87	AAAdA	2
466.56	473.69	515.97	AMdA	4
466.77	473.90	516.18	MAdM	10
467.93	474.24	514.60	MMA	7
467.27	474.40	516.68	MMdM	12
467.41	474.54	516.82	MMdA	8
467.47	474.60	516.88	MAdA	6
469.58	475.89	516.25	MAA	5
479.85	486.16	526.52	MAM	9
492.64	498.94	539.30	MMM	11

Ao observar a tabela 6 comparando as métricas de decisão o modelo a ser considerado é o modelo 3 ou ETS(AMA), desse modo o componente de erro é aditivo, o de tendência multiplicativo e o de sazonalidade é aditivo.

2.4 Análise de Resíduos ETS

Assim como no caso dos modelos Arima é preciso realizar análise de resíduos para os modelos ETS, portanto foi feito os gráficos apresentados na Figura 7 e 8 além dos já citados testes KPSS, Ljung-Box e Shapiro-Wilk para estacionaridade, correlação de resíduos e normalidade respectivamente.

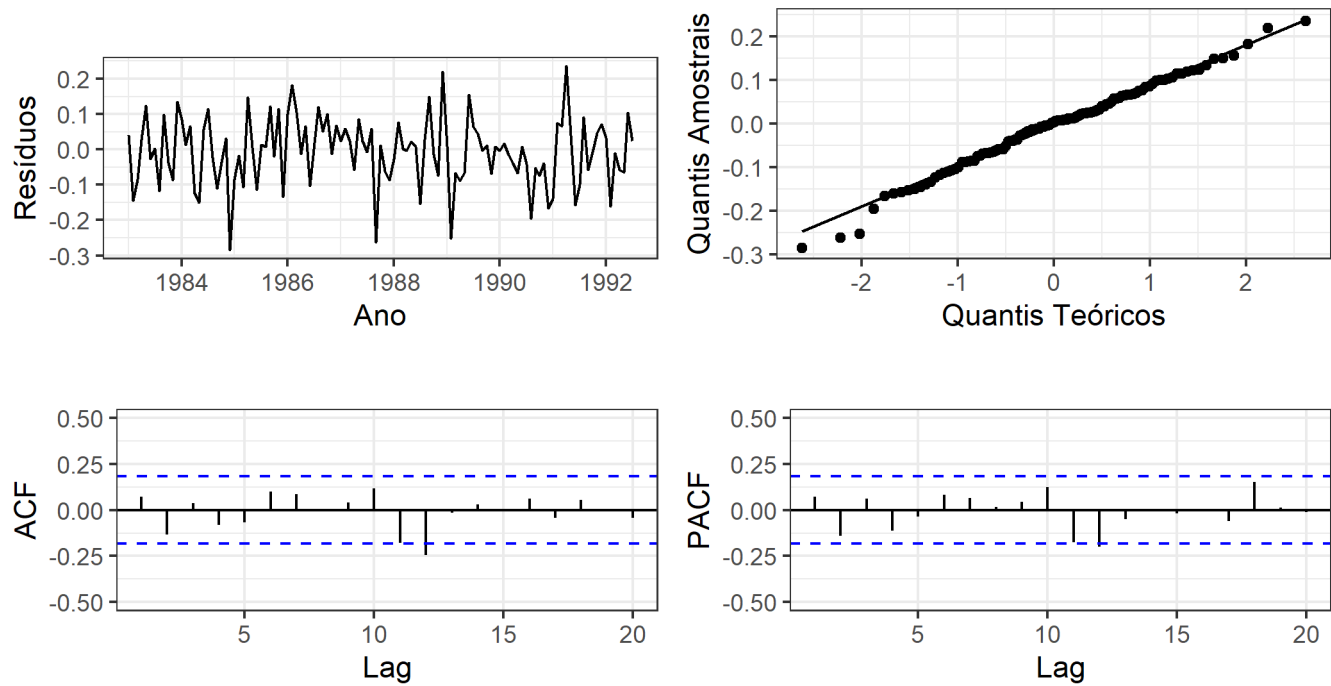


Figura 7: Gráficos de resíduos modelo $ETS(M, Ad, M)$

Tabela 7: Resultado dos teste para modelo $ETS(M, Ad, M)$

Teste	P.Valor
<i>KPSS</i>	0,10
<i>Ljung – Box</i>	0,33
<i>Shapiro – Wilk</i>	0,54

Considerando a Figura 7 o processo apresenta características de estacionaridade, valores ao redor de zero e também variância homogenia, além disso, pode-se verificar a estacionaridade ao observar para o teste KPSS na Tabela 7. No próximo gráfico apresenta os quantis teóricos do resíduo avaliando o ajuste de uma normal, dado o comportamento de praticamente todos os pontos estarem sob a reta, parece razoável afirmar normalidade dos resíduos. Isso hipótese se torna ainda mais evidente quando o resultado do teste de shapiro-wilk apresenta um p.valor maior que 0,1. Por fim, para verificar a independência entre os resíduos apresenta-se os gráficos ACF e PACF com praticamente nenhuma auto-correlação significativa em nenhum momento e de fato quando observa o teste Ljung-Box com p.valor de 0,33 não há motivos para rejeitar a hipótese nula, logo a afirmação de independência parece mais razoável.

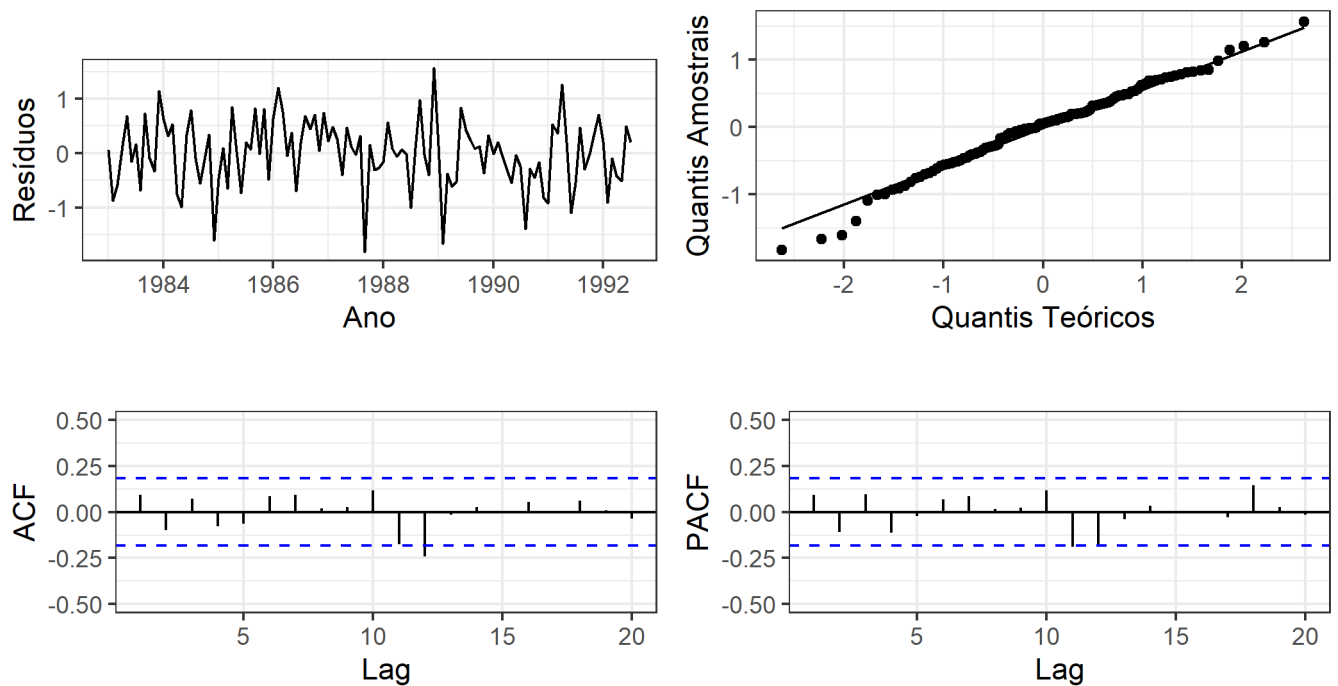


Figura 8: Gráficos de resíduos modelo transformação de Box-Cox $ETS(A, M, A)$

Tabela 8: Resultado dos teste para modelo $SARIMA(0, 1, 1) \times (1, 0, 2)_{12}$

Teste	P.Valor
$KPSS$	0,10
$Ljung - Box$	0,40
$Shapiro - Wilk$	0,29

Ao observar a Figura 8 apresenta-se novamente uma série de resíduos no tempo com características de estacionaridade como a média ao redor de zero e variância constante, o teste KPSS com resultado apresentando na Tabela 8. Outro gráfico apresentado é o de quantis teóricos da normal, com muitos pontos sob a reta a afirmação de normalidade parece razoável isso pode ser verificado pelo teste de shapiro-wilk apresentado em 8 que apresenta um p.valor de 0,29, ou seja os resíduos seguem uma normal. E por fim o teste de Jjung-Box e os gráficos de autocorrelação não apresentam razões para questionar a independência dos resíduos.

3 Resultados

Nessa seção é apresentado os resultados das previsões com 4 dos modelos selecionados. O modelo SARIMA selecionado manualmente, o modelo SARIMA selecionado com a transformação de Box-Cox, modelo ETS(M,Ad,M) sem transformação e o modelo ETS(A,M,A) com transformação de Box-Cox.

3.1 Desempenho preditivo por janelas deslizantes

Uma análise possível para avaliar a capacidade preditiva do modelo é a avaliação de janelas deslizantes, na qual são consideradas múltiplas janelas de treinamento e de teste, cada horizonte de previsão é avaliado múltiplas vezes.

Tabela 9: Resultados da média absoluta do erro para diferentes horizontes

	MAE_ets_cox	MAE_ets	MAE_arima	MAE_arima_cox
h=1	692.16	709.17	540.89	408.83
h=2	548.54	541.97	506.11	446.28
h=3	573.48	568.56	500.88	472.78
h=4	879.07	839.24	569.83	482.68
h=5	796.21	683.93	638.06	460.79

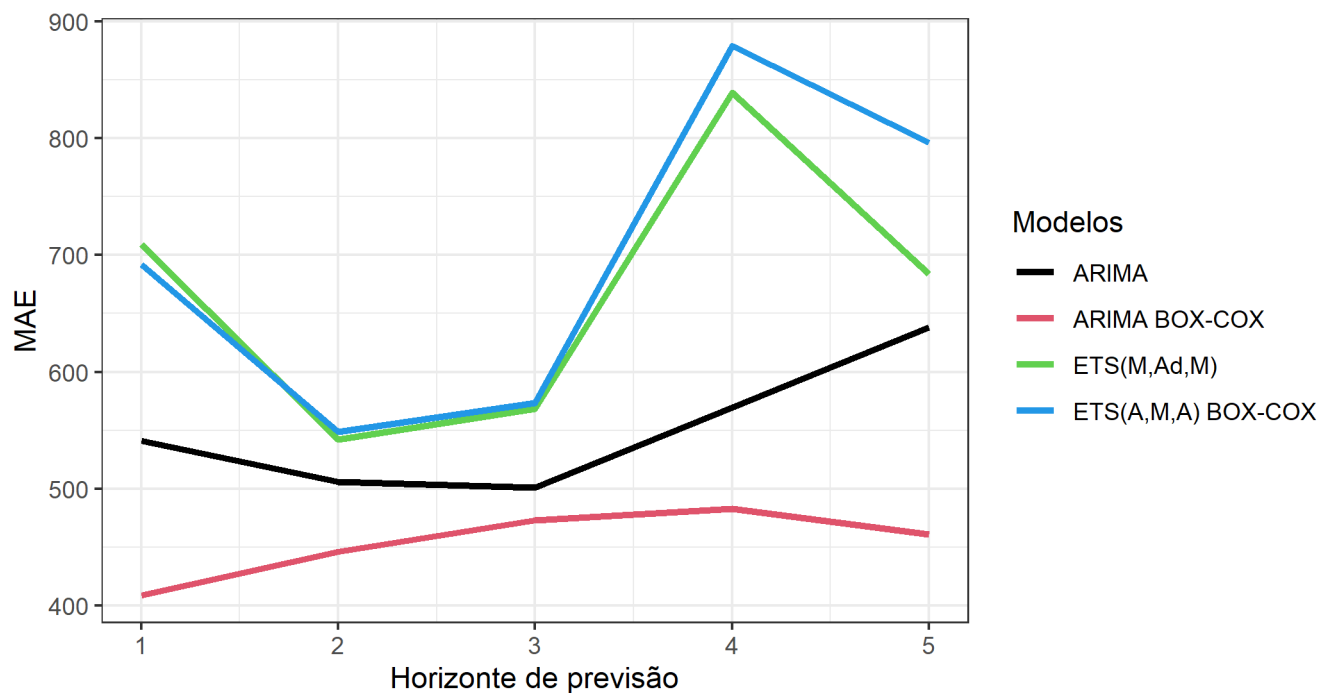


Figura 9: Média do erro absoluto dos modelos selecionados

Fica claro com os valores apresentados na Tabela 9 que os modelos arima apresentam para todos horizontes de previsão valores de erro médio absoluto menores que os modelos

ets. Entre eles o de menor erro médio absoluto é o modelo $SARIMA(0, 1, 1) \times (1, 0, 2)_{12}$ com transformação de Box-cox. Isso fica ainda mais evidente quando observa a representação gráfica da Tabela 9 na Figura ?? no qual fica claro a discrepância dos valores do modelo SARIMA com transformação de Box-cox. Sendo assim, nessa avaliação de desempenho preditivo o modelo $SARIMA(0, 1, 1) \times (1, 0, 2)_{12}$ com transformação de Box-cox é melhor que o selecionado manualmente, porém esse mesmo modelo aplicado a série original possuía um AIC maior.

3.2 Previsões pontuais e intervalares

Mesmo utilizando as janelas deslizantes outra maneira de definir o poder preditivo dos modelos é considerando o erro de previsão de dados inéditos com intervalos de confiança de 95%. No caso da série de teste como é um caso teórico já foi separado de antemão 18 observações que fazem parte do mesmo processo que não foram incluídas em nenhum momento do trabalho até então.

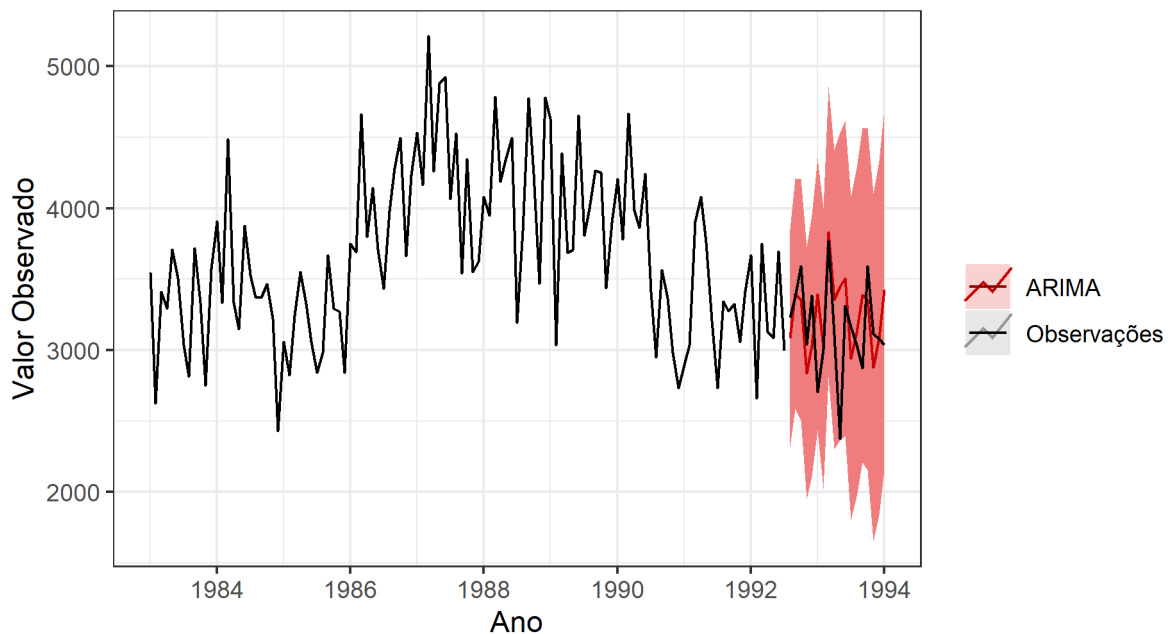


Figura 10: Predições modelo $SARIMA(0, 1, 1) \times (1, 0, 2)_{12}$

Ao observar a Figura 10 as previsões parecem coerentes uma vez que os dados observados estão na cobertura do intervalo de confiança. Além disso a série feita pelas previsões pontuais parece acompanhar boa parte dos casos de observações.

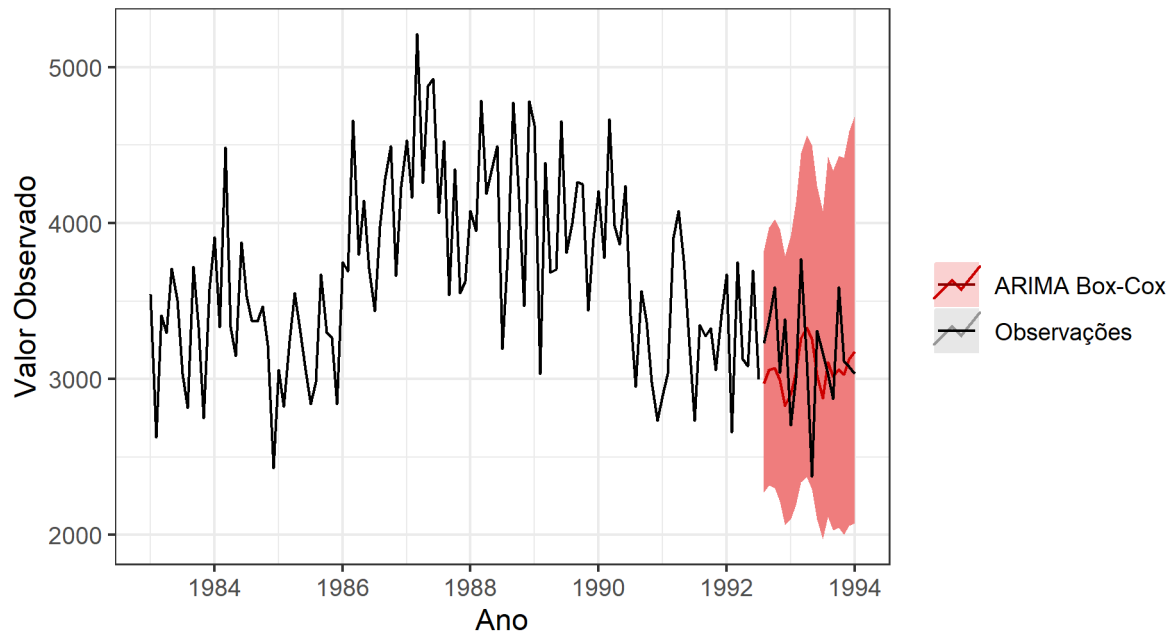


Figura 11: Predições modelo com transformação de Box-cox $SARIMA(0, 1, 1) \times (1, 0, 1)_{12}$

Considerando a 11 observa-se que mesmo o caso das observações estarem compreendidas no intervalo de confiança de 95%, o caso das previsões pontuais parece não acompanhar as observações de fato, os picos ficam distantes e não parece ter captado essa variância.

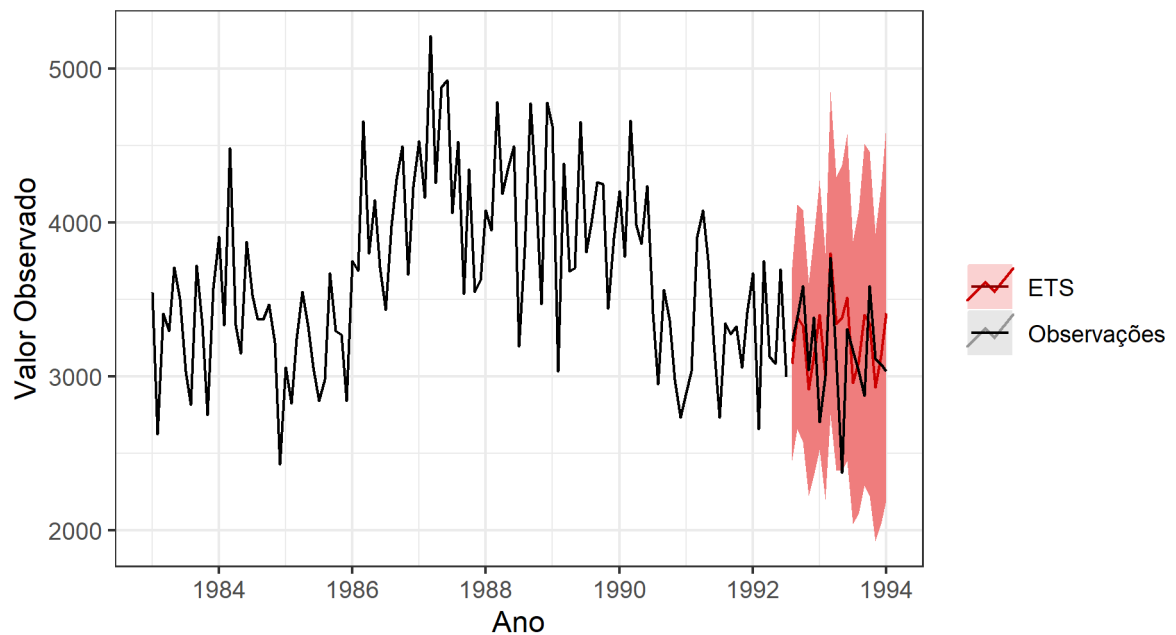


Figura 12: Predições modelo $ETS(M, Ad, M)$

Considerando o modelo ETS sem nenhuma transformação na Figura 12 observa-se um comportamento similar ao que ocorreu na 10 com todos os valores observados compre-

endidos pelos limites do intervalo confiança e a série de previsões pontuais acompanhando parte do movimento dos valores reais.

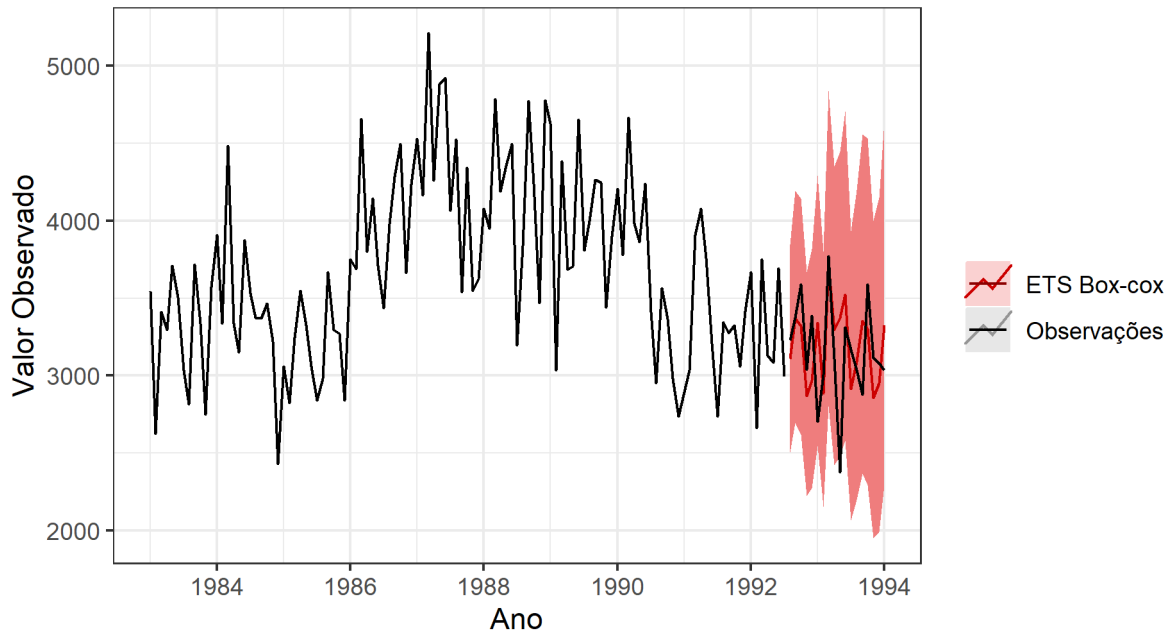


Figura 13: Predições modelo com transformação de Box-Cox $ETS(A, M, A)$

Ao observar a Figura 13 nota-se comportamento similar ao da Figura 12 porém os intervalos de confiança um pouco mais estreitos, deixando um valor mínimo da série observada de fora.

3.2.1 Benchmarks

Tabela 10: Benchmarks

modelos	mae	mape
arima	273.38	9.45
arima_cox_box	285.24	9.15
ets	257.40	8.91
ets_cox_box	274.18	9.39
autoarima	284.00	9.12
ses	268.84	9.04
holt	267.44	8.98
autoets	267.32	9.18
stlf	288.65	9.90
bats	281.40	9.68
tbats	271.11	9.32

Considerando o Benchmarking com diversos outros modelos que estão introduzidos dentro da biblioteca *Forecast* no Software R, considerando a Média Absoluta do Erro e

Média Absoluta do Erro Proporcional, respectivamente na Tabela 10, MAE e MAPE. Essa métrica por ser uma média do erro entre o valor predito e o observado, deve-se considerar que quanto menor, maior capacidade preditiva do modelo, considerando isso, o modelo que teve o melhor resultado foi o ETS(M,Ad,M) sem nenhuma transformação.

4 Conclusão

Mesmo tratando de um trabalho com dados não reais, ou seja, com propósito de aprendizado, decidir um modelo eficaz e correto não é uma tarefa simples. Uma vez que foi proposto 4 modelos, todos tiveram análise de resíduos que confirmaram os pressupostos, ou seja, não houve o descarte de nenhum modelo nessa etapa. Na medida de poder preditivo por janelas deslizantes o modelo com menor erro foi o Sarima com transformação de box-cox, porém quando o caso foi para observações não trabalhadas ainda, o modelo apresentou graficamente e por métricas o pior resultado. Por outro lado, o melhor resultado no caso de novas observações foi o do modelo ETS sem transformação, porém, esse teve um resultado inferior quando se tratou das janelas deslizantes. Dessa maneira, o modelo que parece ter trazido bons resultados em ambas ocasiões e foi consistente durante toda a análise é o modelo Sarima sem transformação. Sendo, assim eu não descartaria nenhum dos outros dois modelos mas escolheria esse Sarima $SARIMA(0, 1, 1) \times (1, 0, 2)_{12}$ como modelo final.

A Tabela de Previsões Pontuais e Intervalares

Tabela 11: Predições modelo com transformação de Box-cox $SARIMA(0, 1, 1) \times (1, 0, 1)_{12}$

Time	Previsão Pontual	Limite Inferior 95%	Limite Superior 95%
Aug 1992	2970.48	2272.24	3825.23
Sep 1992	3059.14	2318.24	3972.35
Oct 1992	3071.50	2303.28	4025.68
Nov 1992	2991.02	2215.65	3962.69
Dec 1992	2826.40	2064.14	3791.56
Jan 1993	2901.07	2101.66	3919.18
Feb 1993	3041.82	2190.30	4131.06
Mar 1993	3263.01	2340.42	4446.47
Apr 1993	3329.12	2370.31	4565.61
May 1993	3254.37	2292.80	4503.83
Jun 1993	3025.40	2100.45	4239.98
Jul 1993	2876.88	1971.71	4076.51
Aug 1993	3107.89	2117.27	4426.42
Sep 1993	3015.60	2029.93	4338.77
Oct 1993	3061.19	2044.83	4433.03
Nov 1993	3027.91	2002.58	4421.62
Dec 1993	3128.35	2057.54	4589.62
Jan 1994	3178.94	2076.36	4690.95

Tabela 12: Predições modelo $SARIMA(0, 1, 1) \times (1, 0, 2)_{12}$

Time	Previsão Pontual	Limite Inferior 95%	Limite Superior 95%
Aug 1992	3087.04	2318.53	3855.54
Sep 1992	3396.87	2587.05	4206.69
Oct 1992	3358.20	2509.07	4207.32
Nov 1992	2838.82	1952.13	3725.51
Dec 1992	3031.77	2109.04	3954.49
Jan 1993	3393.82	2437.43	4350.21
Feb 1993	3013.99	2024.39	4003.60
Mar 1993	3833.40	2811.66	4855.14
Apr 1993	3357.93	2305.03	4410.82
May 1993	3447.73	2364.59	4530.88
Jun 1993	3506.67	2394.09	4619.25
Jul 1993	2943.04	1801.78	4084.30
Aug 1993	3126.27	1969.27	4283.28
Sep 1993	3387.87	2207.64	4568.10
Oct 1993	3357.85	2154.84	4560.86
Nov 1993	2878.50	1653.13	4103.87
Dec 1993	3085.87	1838.55	4333.20
Jan 1994	3425.64	2156.49	4694.78

Tabela 13: Predições modelo ETS(M,Ad,M)

Time	Previsão Pontual	Limite Inferior 95%	Limite Superior 95%
Aug 1992	3082.15	2455.71	3708.59
Sep 1992	3388.05	2659.07	4117.03
Oct 1992	3331.04	2576.68	4085.40
Nov 1992	2917.37	2225.25	3609.48
Dec 1992	3127.44	2353.22	3901.67
Jan 1993	3400.66	2525.08	4276.25
Feb 1993	2998.96	2198.14	3799.79
Mar 1993	3799.44	2749.76	4849.13
Apr 1993	3341.41	2388.36	4294.47
May 1993	3382.70	2388.47	4376.92
Jun 1993	3512.20	2450.23	4574.17
Jul 1993	2960.69	2041.11	3880.27
Aug 1993	3095.03	2108.86	4081.20
Sep 1993	3401.85	2291.24	4512.45
Oct 1993	3344.25	2226.81	4461.70
Nov 1993	2928.65	1928.09	3929.20
Dec 1993	3139.22	2043.64	4234.81
Jan 1994	3413.14	2197.35	4628.94

Tabela 14: Predições modelo com transformação de Box-Cox ETS(A,M,A)

Time	Previsão Pontual	Limite Inferior 95%	Limite Superior 95%
Aug 1992	3108.49	2502.31	3844.53
Sep 1992	3368.09	2693.58	4195.72
Oct 1992	3319.68	2619.09	4145.77
Nov 1992	2870.38	2222.22	3669.02
Dec 1992	2967.89	2270.71	3819.05
Jan 1993	3341.54	2544.83	4292.59
Feb 1993	2885.95	2152.35	3799.56
Mar 1993	3721.40	2802.75	4838.74
Apr 1993	3296.82	2421.29	4352.46
May 1993	3373.37	2478.94	4448.83
Jun 1993	3521.24	2577.16	4709.55
Jul 1993	2913.44	2066.89	3935.04
Aug 1993	3092.95	2199.59	4199.95
Sep 1993	3351.55	2368.68	4556.86
Oct 1993	3303.33	2293.70	4535.86
Nov 1993	2855.77	1948.07	4000.57
Dec 1993	2952.90	1989.67	4147.44
Jan 1994	3325.11	2294.97	4642.84

B Códigos em R

```
library(Mcomp)
library(forecast)
library(tidyverse)
library(tseries)
library(xtable)
library(cowplot)
library(rstudioapi)

# Trabalho de S r i e s – ARIMA – ID 1211
setwd(dirname(getActiveDocumentContext()$path))
getwd()
# Pacotes

data(M3)

# ID – 1936
id <- 1936
M3[[id]]
h <- M3[[id]]$h
treino <- M3[[id]]$x
teste <- M3[[id]]$xx

M3[[id]]$period
M3[[id]]$st
M3[[id]]$type
M3[[id]]$n
# Apresenta o da S r i e
autoplot(treino) +
  labs(x = "Ano", y = "Valor Observado") +
  theme_bw()+
  autolayer(teste, series="Teste") +
  scale_colour_manual(values = c("Teste" = "#A01400"), breaks = c("Teste"

ggsave("imagens//descricao-original.png", width = 158, height = 93, units

lambda <- treino %>% BoxCox.lambda()

treino_box <- treino %>% BoxCox(lambda = lambda)

autoplot(treino_box) +
  labs(x = "Ano", y = "Valor Observado Transformado") +
  theme_bw()

# Decomposi o
decomp.stl <- stl(treino, s.window = 13)
autoplot(decomp.stl) +
  labs(x = "Ano") +
```

```

  theme_bw()
  ggsave("imagens/decomposicao-stl.png", width = 158, height = 93, units = "mm")

# Teste de estacionaridade
kpss.test(treino)

# Teste de sazonalidade
ocsb.test(treino)

# Seleção manual do modelo ARIMA
ndiffs(treino) # ndiffs(diff(treino, lag = 4)) ## precisa da segunda diferen
nsdiffs(treino) # nsdiffs(diff(treino))

ndiffs(treino_box) # ndiffs(diff(treino, lag = 4)) ## precisa da segunda dife
nsdiffs(treino_box) # nsdiffs(diff(treino))

cbind('Série Original' = treino, 'Primeiras Diferenças' = diff(treino), 'Pr
  autoplot(facets = T) +
  labs(x = "Ano", y = "Valor Observado") +
  theme_bw() +
  ggsave("Graficos/selecao-diff.png", width = 158, height = 93, units = "mm")

d_treino <- diff(treino) # série estacionária seria a d_treino com apenas u
# d_treino <- treino
#w <- diff(treino) %>% diff(lag=12)
ndiffs(d_treino)
kpss.test(d_treino) #### depois de tirar a diferença estacionário

nsdiffs(d_treino)

ocsb.test(d_treino)

par(mfrow=c(1,1))

plot_acf<- ggAcf(d_treino, lag.max = 12*3) +
  scale_y_continuous(limits = c(-1, 1)) +
  scale_x_continuous(breaks = seq(0,36,6))+
  labs(title = "") +
  theme_bw()

plot_pacf<- ggPacf(d_treino, lag.max = 12*3) +
  scale_y_continuous(limits = c(-1, 1)) +
  scale_x_continuous(breaks = seq(0,36,6))+
  labs(title = "") +
  theme_bw()

```

```
ggAcf(treino, lag.max = 12*3) +
  scale_y_continuous(limits = c(-1, 1)) +
  scale_x_continuous(breaks = c(1:36))+
  labs(title = "") +
  theme_bw()
```

```
ggPacf(treino, lag.max = 12*3) +
  scale_y_continuous(limits = c(-1, 1)) +
  scale_x_continuous(breaks = c(1:36))+
  labs(title = "") +
  theme_bw()
```

```
plot_grid(plot_acf, plot_pacf, ncol = 2)
ggsave("imagens/acf_pacf.png", width = 178, height = 93, units = "mm")
```

```
# Sabemos que d=1 e D=0
#
# Demais argumentos:
# p = 2 ou 1
# q = 2 ou 1
# P = 0
# Q = 2 ou 1
```

```
## Modelo candidato para {x}
# —> SARIMA (2,1,2) x (0,0,0)
```

```
### testador de modelos hahaha
melhor_AICc = Inf
for(p in 0:2){
  for(q in 0:2){
    for(Q in 1:2){
      fit = Arima(treino, order=c(p,1,q), seasonal=c(1,0,Q))
      if(fit$aicc < melhor_AICc){
        melhor_AICc = fit$aicc
        cat("p = ", p, ", q = ", q, ", P = ", P, ", Q = ", Q, ", AICc =", f
      }
    }
  }
}
```

```
#Arima(treino, order=c(0,1,2), seasonal=c(1,0,2))
#Arima(treino, order=c(0,1,2), seasonal=c(1,0,1))
```

```
mod_arima <- Arima(treino, order = c(0, 1, 1), seasonal = c(1,0,2))
mod_arima

mode_auto_arima <- auto.arima(treino)

res1 <- residuals(mod_arima)

res_auto <- residuals( mode_auto_arima)

mod_box_cox <- auto.arima(treino, lambda = lambada)

res_cox <- residuals(mod_box_cox)

shapiro.test(res_cox)

kpss.test(res1)

Box.test(res1, lag=20, type = 'Ljung-Box')

shapiro.test(res_auto)

kpss.test(res_auto)

Box.test(res_auto, lag=20, type = 'Ljung-Box', fitdf=4)

##### An lise de res duos do ARIMA #####

### Manual

res1 <- residuals(mod_arima)

# Estacionaridade
residual <- autoplot(res1) +
  labs(x = "Ano", y = "Res duos") +
  theme_bw()

kpss.test(res1)

# Independência
ACF_res <- ggAcf(res1, lag.max = 4*5) +
  scale_y_continuous(limits = c(-.5, .5)) +
  labs(title = "") + theme_bw()

PACF_res <- ggPacf(res1, lag.max = 4*5) +
  scale_y_continuous(limits = c(-.5, .5)) +
```



```

      labs(title = "") + theme_bw()

Box.test(res1, lag = 20, type = "Ljung-Box")

# Normalidade
QQplot_res <- data.frame(res1) %>%
  ggplot(aes(sample = res1)) +
  stat_qq() +
  stat_qq_line() + labs(x = "Quantis Teóricos", y = "Quantis

shapiro.test(res1)

plot_grid(residual, QQplot_res, ACF_res, PACF_res, nrow = 2)
ggsave("imagens/residuos_arima.png", width = 178, height = 93, units = "m

#### Auto-arima

res_auto <- residuals(mode_auto_arima)

# Estacionaridade
residual <- autoplot(res_auto) +
  labs(x = "Ano", y = "Res duos") +
  theme_bw()

kpss.test(res_auto)

# Independência
ACF_res <- ggAcf(res_auto, lag.max = 4*5) +
  scale_y_continuous(limits = c(-.5, .5)) +
  labs(title = "") + theme_bw()

PACF_res <- ggPacf(res_auto, lag.max = 4*5) +
  scale_y_continuous(limits = c(-.5, .5)) +
  labs(title = "") + theme_bw()

Box.test(res_auto, lag = 20, type = "Ljung-Box")

# Normalidade
QQplot_res <- data.frame(res_auto) %>%
  ggplot(aes(sample = res_auto)) +
  stat_qq() +

```

```

stat_qq_line() + labs(x = "Quantis Teóricos", y = "Quantis Amostrais") + t

shapiro.test(res_auto)

plot_grid(residual, QQplot_res, ACF_res, PACF_res, nrow = 2)
ggsave("imagens/residuos_auto.png", width = 178, height = 93, units = "mm")

#### Transformada Box-Cox

res_cox <- residuals(mod_box_cox)

# Estacionaridade
residual <- autoplot(res_cox) +
  labs(x = "Ano", y = "Resíduos") +
  theme_bw()

kpss.test(res_cox)

# Independência
ACF_res <- ggAcf(res_cox, lag.max = 4*5) +
  scale_y_continuous(limits = c(-.5, .5)) +
  labs(title = "") + theme_bw()

PACF_res <- ggPacf(res_cox, lag.max = 4*5) +
  scale_y_continuous(limits = c(-.5, .5)) +
  labs(title = "") + theme_bw()

Box.test(res_cox, lag = 20, type = "Ljung-Box")

# Normalidade
QQplot_res <- data.frame(res_cox) %>%
  ggplot(aes(sample = res_cox)) +
  stat_qq() +
  stat_qq_line() + labs(x = "Quantis Teóricos", y = "Quantis Amostrais") + t

shapiro.test(res_cox)

plot_grid(residual, QQplot_res, ACF_res, PACF_res, nrow = 2)
ggsave("imagens/residuos_cox.png", width = 178, height = 93, units = "mm")

##### Seleção manual do modelo ETS #####

ets1 <- ets(treino, model = "AAA", damped = FALSE, restrict = FALSE) # ETS(A,

```

```

ets2 <- ets(treino , model = "AAA", damped = TRUE, restrict = FALSE)
# ETS(A,Ad,A)
ets3 <- ets(treino , model = "AMA", damped = FALSE, restrict = FALSE) # ET
ets4 <- ets(treino , model = "AMA", damped = TRUE, restrict = FALSE)
# ETS(A,Md,A)
ets5 <- ets(treino , model = "MAA", damped = FALSE, restrict = FALSE) # ET
ets6 <- ets(treino , model = "MAA", damped = TRUE, restrict = FALSE)
# ETS(M,Ad,A)
ets7 <- ets(treino , model = "MMA", damped = FALSE, restrict = FALSE) # ET
ets8 <- ets(treino , model = "MMA", damped = TRUE, restrict = FALSE)
# ETS(M,Md,A)
ets9 <- ets(treino , model = "MAM", damped = FALSE, restrict = FALSE) # ET
ets10 <- ets(treino , model = "MAM", damped = TRUE, restrict = FALSE)
# ETS(M,Ad,M)
ets11 <- ets(treino , model = "MMM", damped = FALSE, restrict = FALSE) # E
ets12 <- ets(treino , model = "MMM", damped = TRUE, restrict = FALSE)
# ETS(M,Md,M)

aics<- rbind(ets1$aic ,ets2$aic ,ets3$aic ,ets4$aic ,ets5$aic ,ets6$aic ,
             ets7$aic ,ets8$aic ,ets9$aic ,ets10$aic ,ets11$aic ,ets12$aic)

bics<- rbind(ets1$bic ,ets2$bic ,ets3$bic ,ets4$bic ,ets5$bic ,ets6$bic ,
             ets7$bic ,ets8$bic ,ets9$bic ,ets10$bic ,ets11$bic ,ets12$bic)

aiccs<- rbind(ets1$aicc ,ets2$aicc ,ets3$aicc ,ets4$aicc ,ets5$aicc ,ets6$aicc ,
             ets7$aicc ,ets8$aicc ,ets9$aicc ,ets10$aicc ,ets11$aicc ,ets12$aicc)

Criterios_Ets <- data.frame(aics ,aiccs ,bics)
Criterios_Ets$modelos <- c("AAA","AAdA","AMA","AMdA","MAA","MAdA","MMA","MAM")
Criterios_Ets$mods <- c(1:12)
Criterios_Ets <- Criterios_Ets %>% arrange(aiccs)
Criterios_Ets <- Criterios_Ets[, -5]
Criterios_Ets <- Criterios_Ets[1:4,]

xtable(Criterios_Ets)

##### transforma o

ets1 <- ets(treino , model = "AAA", damped = FALSE, restrict = FALSE, lambda = 0.01)
ets2 <- ets(treino , model = "AAA", damped = TRUE, restrict = FALSE, lambda = 0.01)
# ETS(A,Ad,A)
ets3 <- ets(treino , model = "AMA", damped = FALSE, restrict = FALSE, lambda = 0.01)
ets4 <- ets(treino , model = "AMA", damped = TRUE, restrict = FALSE, lambda = 0.01)
# ETS(A,Md,A)
ets5 <- ets(treino , model = "MAA", damped = FALSE, restrict = FALSE, lambda = 0.01)
ets6 <- ets(treino , model = "MAA", damped = TRUE, restrict = FALSE, lambda = 0.01)
# ETS(M,Ad,A)

```

```

ets7 <- ets(treino , model = "MMA", damped = FALSE, restrict = FALSE, lambda =
ets8 <- ets(treino , model = "MMA", damped = TRUE, restrict = FALSE, lambda =
# ETS(M,Md,A)
ets9 <- ets(treino , model = "MAM", damped = FALSE, restrict = FALSE, lambda =
ets10 <- ets(treino , model = "MAM", damped = TRUE, restrict = FALSE, lambda =
# ETS(M,Ad,M)
ets11 <- ets(treino , model = "MMM", damped = FALSE, restrict = FALSE, lambda =
ets12 <- ets(treino , model = "MMM", damped = TRUE, restrict = FALSE, lambda =
# ETS(M,Md,M)

aics<- rbind(ets1$aic ,ets2$aic ,ets3$aic ,ets4$aic ,ets5$aic ,ets6$aic ,
             ets7$aic ,ets8$aic ,ets9$aic ,ets10$aic ,ets11$aic ,ets12$aic)

bics<- rbind(ets1$bic ,ets2$bic ,ets3$bic ,ets4$bic ,ets5$bic ,ets6$bic ,
             ets7$bic ,ets8$bic ,ets9$bic ,ets10$bic ,ets11$bic ,ets12$bic)

aiccs<- rbind(ets1$aicc ,ets2$aicc ,ets3$aicc ,ets4$aicc ,ets5$aicc ,ets6$aicc ,
             ets7$aicc ,ets8$aicc ,ets9$aicc ,ets10$aicc ,ets11$aicc ,ets12$aicc)

Criterios_Ets <- data.frame(aics ,aiccs ,bics)
Criterios_Ets$modelos <- c("AAA","AAdA","AMA","AMdA","MAA","MAdA","MMA","MMdA")
Criterios_Ets$mods <- c(1:12)
Criterios_Ets <- Criterios_Ets %>% arrange(aiccs)
xtable(Criterios_Ets)

##### An lise de res duos do ETS #####

mod_ets <- ets(treino , model = "MAM", damped = TRUE, restrict = FALSE) # ETS(M,Md,M)

res_ets <- residuals(mod_ets)

# Estacionaridade
residual <- autoplot(res_ets) +
  labs(x = "Ano", y = "Res duos") +
  theme_bw()

kpss.test(res_ets)

# Independ ncia
ACF_res <- ggAcf(res_ets , lag.max = 4*5) +
  scale_y_continuous(limits = c(-.5, .5)) +
  labs(title = "") + theme_bw()

PACF_res <- ggPacf(res_ets , lag.max = 4*5) +
  scale_y_continuous(limits = c(-.5, .5)) +
  labs(title = "") + theme_bw()

```

```

Box.test(res_ets, lag = 20, type = "Ljung-Box")

# Normalidade
QQplot_res <- data.frame(res_ets) %>%
  ggplot(aes(sample = res_ets)) +
  stat_qq() +
  stat_qq_line() + labs(x = "Quantis Teóricos", y = "Quantis Amostrais")

shapiro.test(res_ets)

plot_grid(residual, QQplot_res, ACF_res, PACF_res, nrow = 2)
ggsave("imagens/residuos_ets.png", width = 178, height = 93, units = "mm")

#### modelo ETS Box-Cox

mod_ets_cox <- ets(treino, model = "AMA", damped = FALSE, restrict = FALSE)

res_ets_cox <- residuals(mod_ets_cox)

# Estacionaridade

residual <- autoplot(res_ets_cox) +
  labs(x = "Ano", y = "Resíduos") +
  theme_bw()

kpss.test(res_ets_cox)

# Independência
ACF_res <- ggAcf(res_ets_cox, lag.max = 4*5) +
  scale_y_continuous(limits = c(-.5, .5)) +
  labs(title = "") + theme_bw()

PACF_res <- ggPacf(res_ets_cox, lag.max = 4*5) +
  scale_y_continuous(limits = c(-.5, .5)) +
  labs(title = "") + theme_bw()

Box.test(res_ets_cox, lag = 20, type = "Ljung-Box")

# Normalidade
QQplot_res <- data.frame(res_ets_cox) %>%
  ggplot(aes(sample = res_ets_cox)) +
  stat_qq() +
  stat_qq_line() + labs(x = "Quantis Teóricos", y = "Quantis Amostrais")

shapiro.test(res_ets_cox)

```

```
plot_grid(residual , QQplot_res , ACF_res , PACF_res , nrow = 2)
ggsave("imagens/residuos_ets_cox.png", width = 178, height = 93, units = "mm")
```

```
##### Janelas Deslizantes #####
```

```
n <- M3[[1936]]$n
```

```
#### ARIMA
```

```
f_arima <- function(y, h){
  fit = arima(y, order = c(0, 1, 1), seasonal = c(1,0,2))
  forecast(fit , h)
}
```

```
CV_arima = tsCV(y=treino , forecastfunction=f_arima , h=5, window = 14)
MAE_arima = CV_arima %>% abs() %>% colMeans(na.rm=T)
```

```
f_arima_cox <- function(y, h){
  fit = auto.arima(y,lambda = lambda)
  forecast(fit , h)
}
```

```
CV_arima_cox = tsCV(y=treino , forecastfunction=f_arima_cox , h=5, window = 14)
MAE_arima_cox = CV_arima_cox %>% abs() %>% colMeans(na.rm=T)
```

```
#### ETS
```

```
f_ets <- function(y, h){
  fit = ets(treino , model = "MAM", damped = TRUE, restrict = FALSE) # ETS(M,A)
  forecast(fit , h)
}
```

```
CV_ets = tsCV(y=treino , forecastfunction=f_ets , h=5, window = 14)
MAE_ets = CV_ets %>% abs() %>% colMeans(na.rm=T)
```

```
f_ets_cox <- function(y, h){
  fit = ets(treino , model = "AMA", damped = FALSE, restrict = FALSE, lambda
  forecast(fit , h)
}
```

```
CV_ets_cox = tsCV(y=treino , forecastfunction=f_ets_cox , h=5, window = 14)
MAE_ets_cox = CV_ets_cox %>% abs() %>% colMeans(na.rm=T)
```

```
H <- c(1,2,3,4,5)
```

```
resultados <- cbind(MAE_ets_cox,MAE_ets, MAE_arima, MAE_arima_cox,H) %>% as.
```

```

df_plot <- resultados %>% pivot_longer(cols = starts_with("MAE"),
                                       names_to = "Modelos",
                                       names_prefix = "MAE_",
                                       values_to = "MAE",
                                       values_drop_na = T)

df_plot %>%
  ggplot(aes(x=H, y=MAE, group=Modelos, color=Modelos)) +
  geom_line(size=1.2) +
  scale_color_manual(labels=c('ARIMA', "ARIMA BOX-COX", "ETS(M,Ad,M)", "E
  xlab('Horizonte de previs o')+
  theme_bw()

ggsave("imagens/MAE_janelas_deslizantes.png", width = 178, height = 93, u

cbind(MAE_ets_cox, MAE_ets, MAE_arima, MAE_arima_cox, H) %>% xtable()

##### Previs es Pontuais e Intervalares #####

arima_preds <- forecast(mod_arima, h = h, level = c(95))
autoplot(treino) + xlab("Ano") + ylab("Valor Observado") +
  theme_bw() +
  autolayer(arima_preds, series="ARIMA") +
  autolayer(teste, series="Observa es") +
  scale_colour_manual(values = c("ARIMA" = "#8B0000", "Observa es" = "
ggsave("imagens/preds_arima.png", width = 158, height = 93, units = "mm")

xtable(arima_preds %>% as.data.frame())

arima_box_preds <- forecast(mod_box_cox, h = h, level = c(95))
autoplot(treino) + xlab("Ano") + ylab("Valor Observado") +
  theme_bw() +
  autolayer(arima_box_preds, series="ARIMA Box-Cox") +
  autolayer(teste, series="Observa es") +
  scale_colour_manual(values = c("ARIMA Box-Cox" = "#8B0000", "Observa
ggsave("imagens/preds_arima_cox.png", width = 158, height = 93, units = "

xtable(arima_box_preds %>% as.data.frame())

ets_preds <- forecast(mod_ets, h = h, level = c(95))
autoplot(treino) + xlab("Ano") + ylab("Valor Observado") +
  theme_bw() +
  autolayer(ets_preds, series="ETS") +
  autolayer(teste, series="Observa es") +

```

```
scale_colour_manual(values = c("ETS" = "#8B0000", "Observa es" = "black"),
ggsave("imagens/preds_ets.png", width = 158, height = 93, units = "mm")
```

```
ets_preds %>% as.data.frame() %>% xtable()
```

```
ets_cox_preds <- forecast(mod_ets_cox, h = h, level = c(95))
autoplot(treino) + xlab("Ano") + ylab("Valor Observado") +
  theme_bw() +
  autolayer(ets_cox_preds, series="ETS Box-cox") +
  autolayer(teste, series="Observa es") +
  scale_colour_manual(values = c("ETS Box-cox" = "#8B0000", "Observa es" =
ggsave("imagens/preds_ets_cox.png", width = 158, height = 93, units = "mm")
```

```
ets_cox_preds %>% as.data.frame() %>% xtable()
# Compara o com a auto.arima()
```

```
autoarima_preds <- forecast(auto.arima(treino), h = h)
ses_preds <- ses(treino, h = h)
holt_preds <- holt(treino, h = h)
autoets_preds <- forecast(ets(treino), h = h)
stlf_preds <- stlf(treino, h = h)
bats_preds <- forecast(bats(treino), h = h)
tbats_preds <- forecast(tbats(treino), h = h)
```

```
lista <- lst(arima_preds, arima_box_preds, ets_preds, ets_cox_preds, autoarima_
mae <- unlist(lapply(lista, function(x) return(mean(abs(teste - x$mean)))))
mape <- unlist(lapply(lista, function(x) return(mean(abs(100*(teste - x$mean)
names(mae) <- sub("_preds", "", names(lista)); names(mape) <- sub("_preds", ""
cbind(mae, mape)
xtable(cbind(mae, mape))
```