

Trabalho Prático 2

Análise de Séries Temporais - 1/2023

Ana Carolina Vianna - 18/0097261

César Augusto Galvão - 19/0011572

Yan Flávio Vianna - 14/0166149

Table of contents

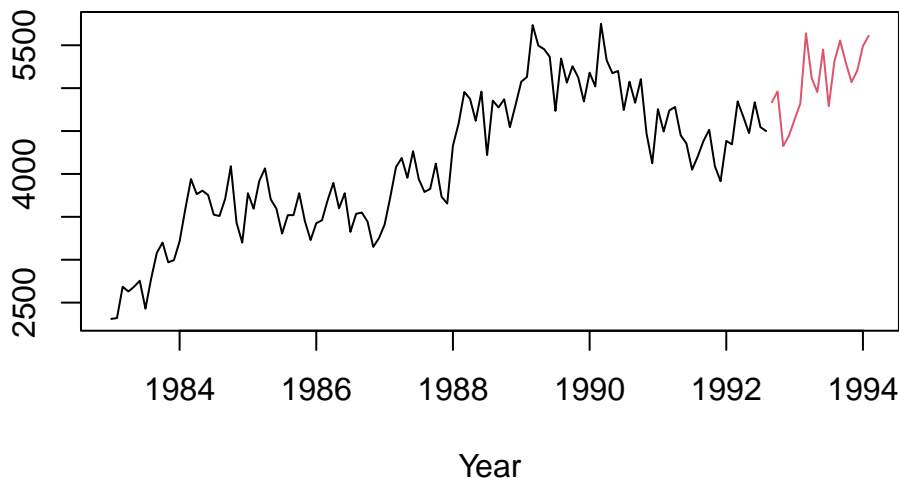
Introdução: série selecionada, características e decomposição	2
Modelos ARIMA: seleção, transformações e resíduos	3
Modelo sem transformação	3
Modelo com transformação	5
Modelos ETS: seleção, transformações e resíduos	7
Modelo sem transformação	7
Modelo com transformação	8
Estudo de desempenho preditivo	10
Resultados da Janela Deslizante	10
Performance em relação aos horizontes de previsão	10
Resultados	10
Apêndice	10

Introdução: série selecionada, características e decomposição

A série temporal escolhida foi a de número *id* correspondente a 2183. De acordo com a definição do próprio pacote, refere-se a *Fluid power shipments - hydraulic index*. Foram realizadas medidas mensais de 1983 a 1992 e o horizonte de previsão requerido é das 18 ocorrências seguintes.

O gráfico da série, com *in* e *out-sample*, é exposto a seguir.

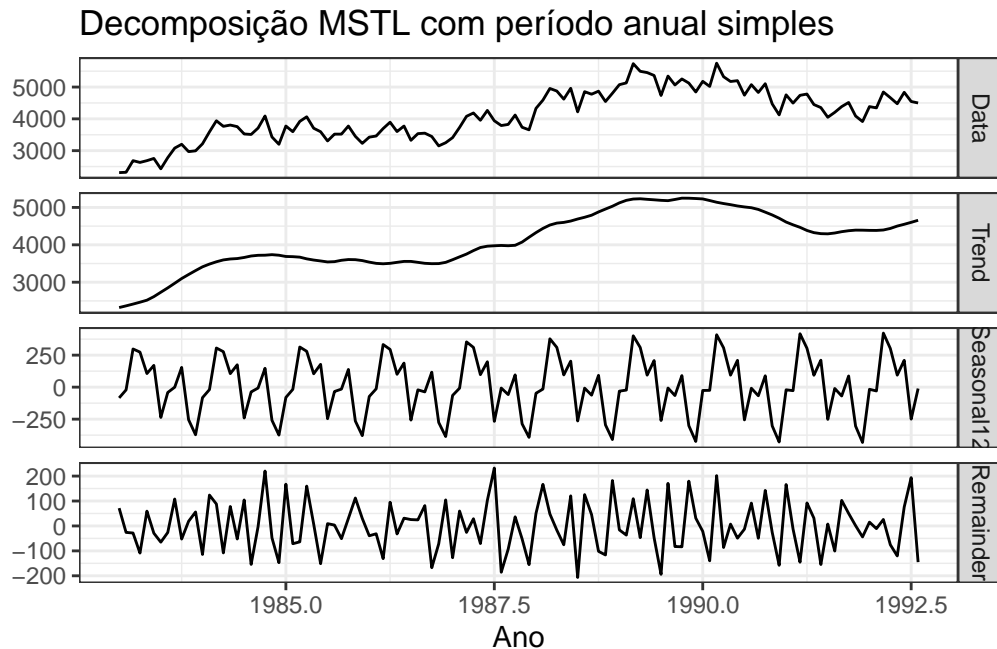
Série Temporal M3-2183



A série aparenta ter dois períodos, pelo menos: um ciclo anual e outro que compreende um período maior. No entanto, ao se tentar decompor a série com múltiplas sazonalidades, obté-se o seguinte:

- **Adicionando uma componente sazonal com ciclo menor que 1 ano** – uma das componentes sazonais apresenta heteroscedasticidade;
- **Adicionando uma componente sazonal com ciclo maior que 1 ano** – resíduos apresentam periodicidade ou heteroscedasticidade.

Optou-se portanto pela decomposição STL (apesar de os dados terem inicialmente formado um objeto `msts`) apenas com a sazonalidade anual, mas fica evidente que esta decomposição não é adequada quando se avalia a componente de tendência, que aparenta ainda carregar algum componente periódico. Os resíduos aparentam um comportamento aleatório e têm média - 0.104, o que é próximo de zero o suficiente considerando a magnitude dos dados da série. A decomposição é exposta a seguir.



Modelos ARIMA: seleção, transformações e resíduos

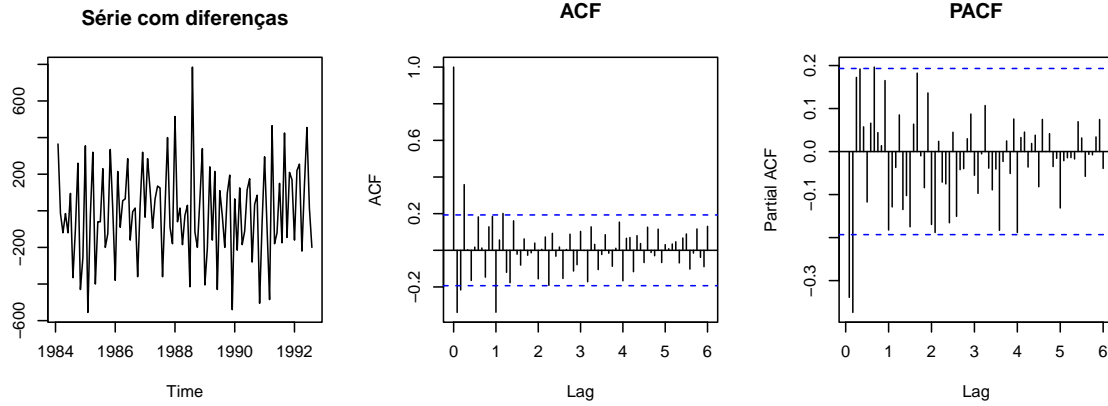
Modelo sem transformação

Seleção

Primeiramente, utilizou-se as funções *ndiffs()* e *nsdiffs()* do pacote *forecast* para identificar quantas diferenças simples e sazonais seriam necessárias para que a série se tornasse estacionária. Concluiu-se pelo resultado dessas funções que são necessárias uma diferenciação simples e uma sazonal. O teste KPSS confirma isso ao não rejeitar a hipótese nula de estacionariedade da série (com diferenças já aplicadas) ao nível de 5% de significância.

	Estatística	p-valor
KPSS Test for Level Stationarity	0.11	0.1

Prosseguimos com a seleção do melhor modelo ARIMA avaliando os gráficos de ACF e PACF. O primeiro parece apresentar quebra no primeiro lag sazonal, enquanto o segundo tem quebra no segundo lag simples. Entretanto, como não fica nítido um comportamento de queda amortizada, preferiu-se utilizar outro critério para a seleção do modelo.

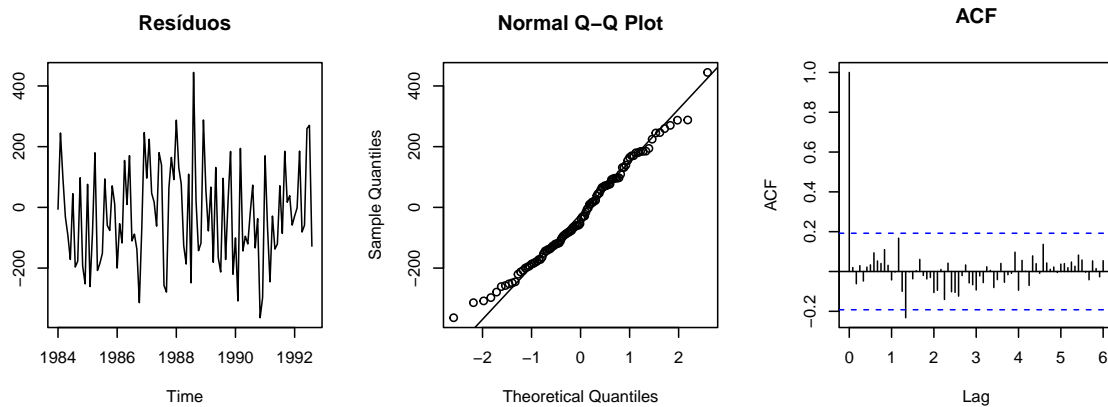


Optou-se pela varredura de combinações de p , q , P e Q , com d e D fixados em 1, como resultado das diferenciações já avaliadas. Utilizando o critério de Akaike corrigido, seleciona-se o modelo $\text{ARIMA}(2, 1, 2) \times (0, 1, 2)_{12}$ para a série, que possui o menor escore entre os modelos testados.

Ao se utilizar a função `auto.arima()`, recebe-se um modelo sugerido $\text{ARIMA}(2, 1, 2) \times (2, 1, 0)_{12}$, porém com AICc superior àquele identificado na varredura. Opta-se pelo modelo selecionado manualmente.

Resíduos

Foram retirados os zeros da inicialização para possibilitar a análise dos resíduos. Observa-se pelo gráfico que os resíduos são aleatórios e aparentemente centrados em zero, com variação constante. Além disso, verifica-se uma distribuição aproximadamente normal, mas com caudas mais pesadas. Finalmente, o gráfico ACF apresenta que a autocorrelação dos resíduos está, em sua grande maioria, dentro da banda de confiança, com exceção de um ponto, que extrapola ligeiramente a margem.



Por fim, realiza-se testes de hipótese para independência e normalidade (o teste KPSS para estacionariedade já foi apresentado) e seus resultados são apresentados na tabela a seguir. De fato, o teste de Shapiro-Wilk não rejeita a normalidade da distribuição dos resíduos apesar de o gráfico QQ apresentar caudas pesadas. Além disso, o teste Ljung-Box com *lag* igual a 15 também não rejeita a independência entre os resíduos e, conseqüentemente, os dados da série.

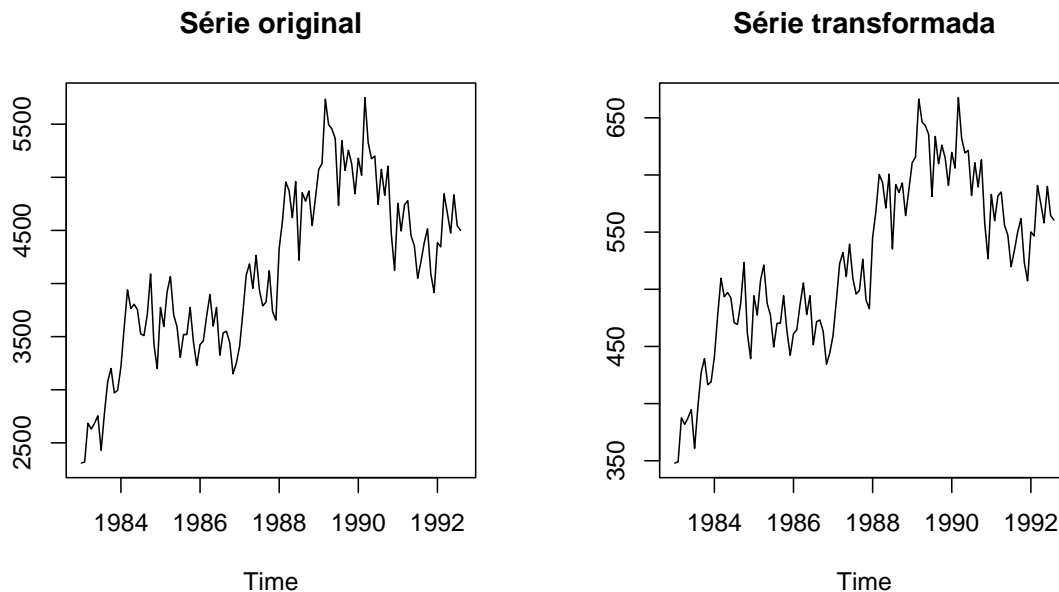
	Estatística	p-valor	Lag
Box-Ljung test	8.90	0.88	15
Shapiro-Wilk normality test	0.99	0.35	

Modelo com transformação

Seleção

Foi utilizada a função *BoxCox.lambda()* do pacote *forecast* para decidir de forma automatizada o melhor valor de lambda para a transformação de Box-Cox. A função sugere um valor de $\lambda = 0.71$.

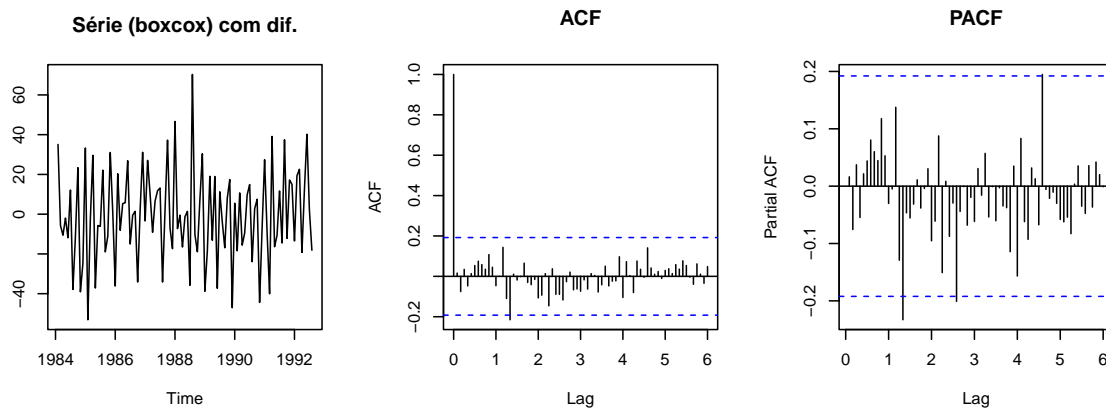
Apesar de haver uma sugestão de transformação, não é possível avaliar graficamente se houve uma diferença significativa no comportamento da série temporal excetuando-se a escala, como se pode ver nos eixos dos gráficos a seguir.



Após aplicar a transformação de Box-Cox na série, utilizou-se as funções `ndiffs()` e `nsdiffs()` para identificar quantas diferenciações simples e sazonais seriam necessárias para que a série se torne estacionária. Concluiu-se que são necessárias uma diferenciações simples e uma diferenciações sazonal, o que é confirmado pelo resultado do teste KPSS nos resíduos da série com as diferenças já aplicadas.

	Estatística	p-valor
KPSS Test for Level Stationarity	0.12	0.1

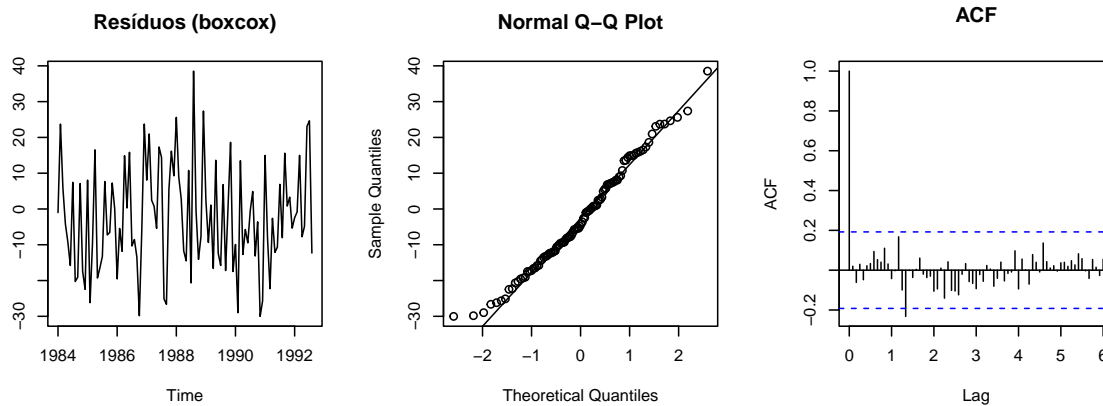
O gráfico da ACF parece apresentar quebra no primeiro lag sazonal, enquanto o PACF tem quebra no segundo lag simples. Entretanto, os gráficos não evidenciam comportamentos claros para a série. Novamente, os resíduos parecem ter média igual a zero.



Foram testadas combinações de p , q , P e Q , com d e D fixados em 1 e, em seguida, selecionou-se o modelo ARIMA que apresentava menor valor do AICc. Temos, então, que o modelo escolhido para a série transformada é um $ARIMA(2, 1, 2) \times (0, 1, 2)_{12}$, assim como no caso da série sem transformação. Utilizando-se a função `auto.arima()` recebe-se uma sugestão de um modelo $ARIMA(3, 1, 1) \times (2, 1, 0)_{12}$ mas, assim como ocorre no modelo sem transformação, opta-se pelo modelo selecionado manualmente por apresentar um AICc menor.

Resíduos

Foram retirados os zeros da inicialização para seguir com a análise dos resíduos. O gráfico da série dos resíduos sugere aleatoriedade e o QQ plot distribuição aproximadamente normal. Por último, o gráfico ACF mostra que a autocorrelação dos resíduos está dentro da banda de confiança, com exceção de um ponto que excede um pouco este limite.



Assim como ocorre para a série não transformada, os testes de Shapiro-Wilk e Ljung-box com lag igual a 15 não apresentam indicação para rejeição de suas hipóteses nulas. Isto é, pode-se dizer que a série transformada tem distribuição normal e seus resíduos são independentes.

	Estatística	p-valor	Lag
Box-Ljung test	8.41	0.91	15
Shapiro-Wilk normality test	0.98	0.23	

Modelos ETS: seleção, transformações e resíduos

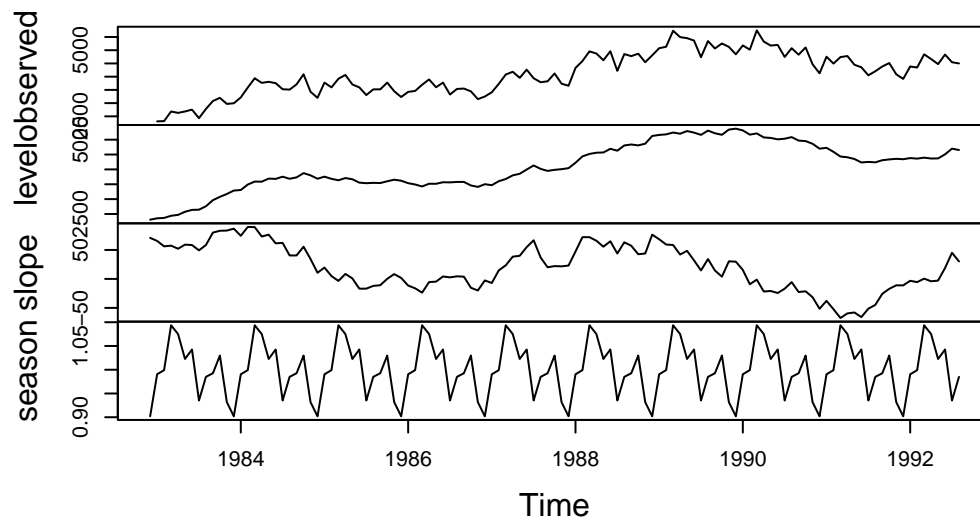
Modelo sem transformação

Seleção

Iniciamos a exploração do modelo

Modelo	AIC	AICc	BIC
ETS(M,Ad,M)	1761.30	1768.36	1810.87
ETS(M,M,M)	1761.94	1769.00	1811.51
ETS(A,Ad,A)	1764.25	1771.30	1813.81
ETS(M,Ad,A)	1767.73	1774.78	1817.29
ETS(M,A,M)	1769.04	1775.29	1815.86
ETS(A,A,A)	1771.20	1777.44	1818.01

Decomposition by ETS(M,Ad,M) method



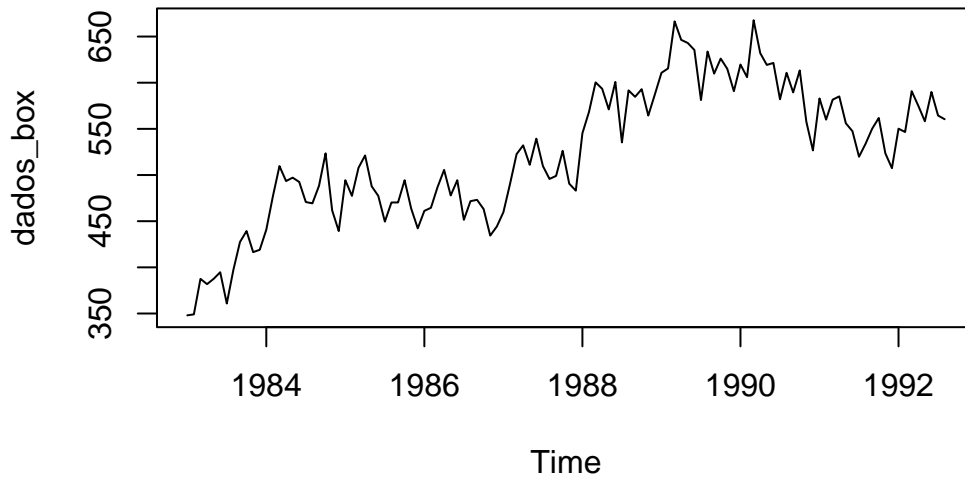
Resíduos

Modelo com transformação

Seleção

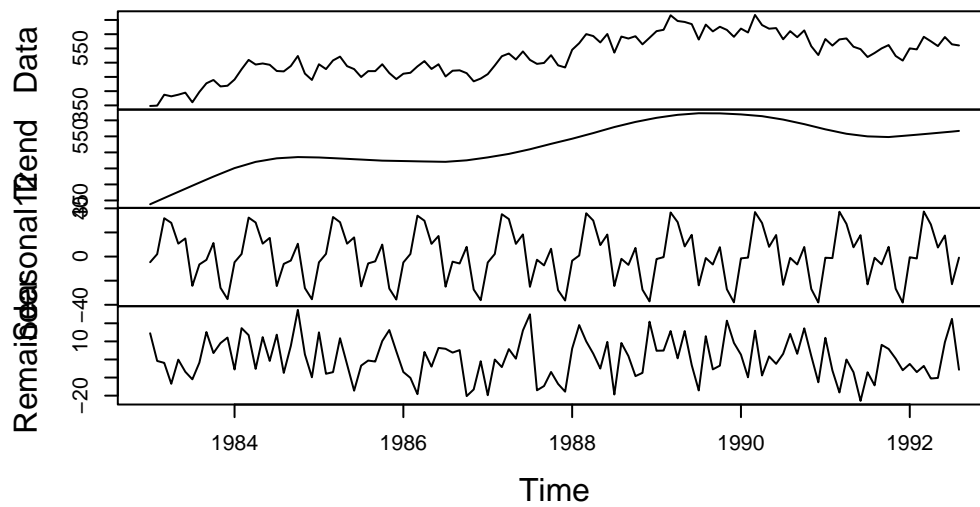
a série com transformacao

Série com transformação Box-Cox $\lambda = 0.712$



decomposicao

Decomposição da série com transformação Box-Cox



selecao do modelo com transformação

Modelo transformado	AIC	AICc	BIC
ETS(M,Ad,M)	1761.30	1768.36	1810.87
ETS(M,M,M)	1761.94	1769.00	1811.51
ETS(Ad,A,A)	1764.25	1771.30	1813.81
ETS(M,Ad,A)	1767.73	1774.78	1817.29
ETS(M,A,M)	1769.04	1775.29	1815.86
ETS(A,A,A)	1771.20	1777.44	1818.01

OS MODELOS SAO OS MESMO, PODEMO SELECIONAR O SEGUNDO MELHOR

Resíduos

Estudo de desempenho preditivo

Resultados da Janela Deslizante

Performance em relação aos horizontes de previsão

ARIMA

ETS

Resultados

apresente em tabelas e gráficos as previsões dos 4 modelos selecionados e também apresente em uma tabela os resultados de acurácia dos 4 modelos selecionados e dos modelos benchmarks. Comente os resultados de modo objetivo;

Apêndice

Todo o projeto de composição deste documento pode ser encontrado aqui: https://github.com/cesar-galvao/trabalhos_series_temporais