



MODELAÇÃO E PREVISÃO DO CONSUMO MENSAL DE ELETRICIDADE EM ESPANHA

Hugo Veríssimo

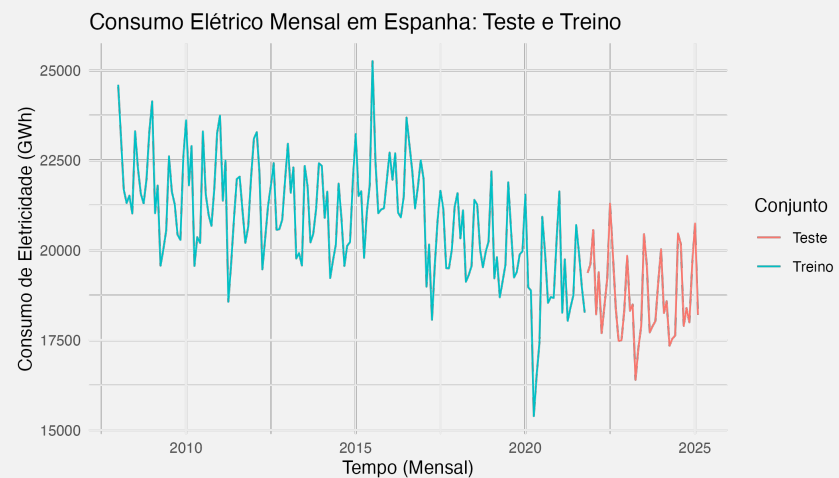
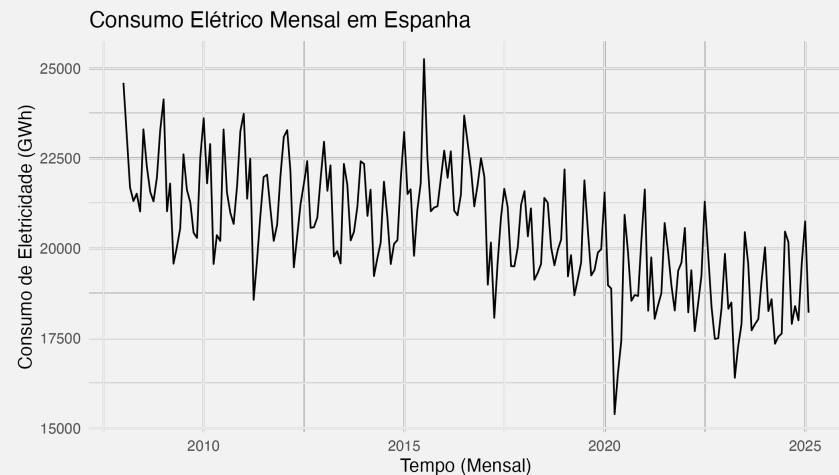
ENQUADRAMENTO

IMPORTÂNCIA DA MONITORIZAÇÃO DO CONSUMO ELÉTRICO

- Papel crucial na gestão energética de países desenvolvidos
- Apoio ao planeamento de capacidade, definição de políticas e avaliação ambiental

RELEVÂNCIA DO CASO DE ESTUDO: ESPANHA

- Economia de grande dimensão energética na Europa
- Contexto ideal para testar a robustez de modelos preditivos em ambientes dinâmicos



ANÁLISE EXPLORATÓRIA

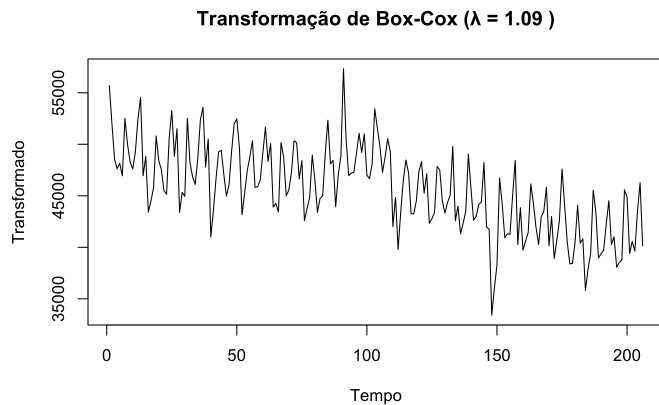
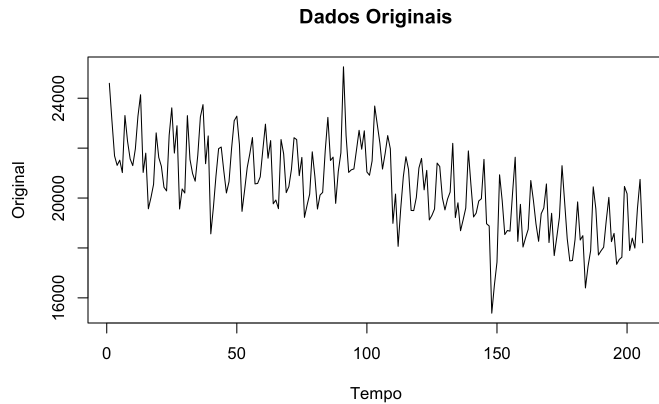
- Consumo mensal de eletricidade (GWh) em Espanha
- Janeiro de 2008 a fevereiro de 2025 (206 observações)
- Sazonalidade? Sim
- Tendência? Talvez
- Variância constante? Talvez
- Outliers? 2020
- Separação em teste e treino (80/20)

ESTACIONARIEDADE

Heterocedasticidade

- Transformação Box-Cox
- ```
studentized Breusch-Pagan test

data: model
BP = 0.35006, df = 1, p-value = 0.5541
Rej. H_0 : Não há evidência de heterocedasticidade
```
- Nenhuma transformação aplicada

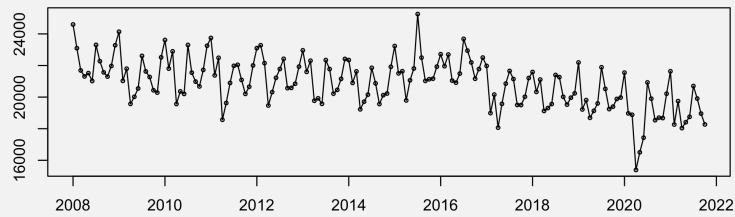




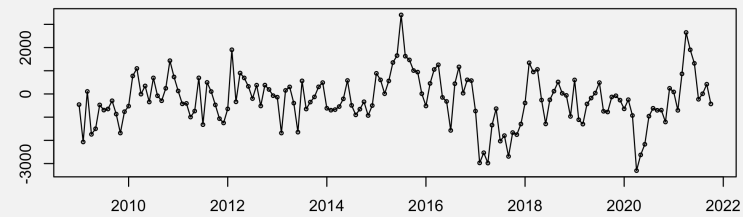
# ESTACIONARIEDADE

## Sazonalidade e Tendência

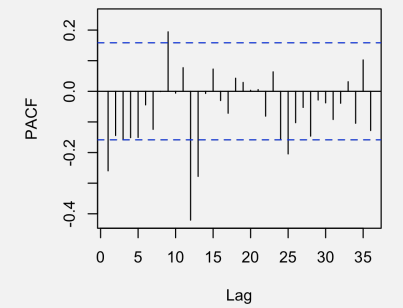
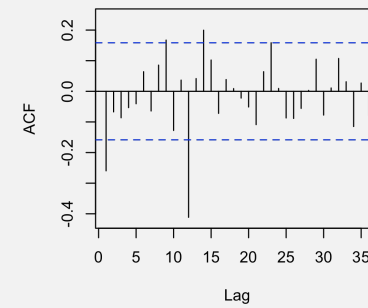
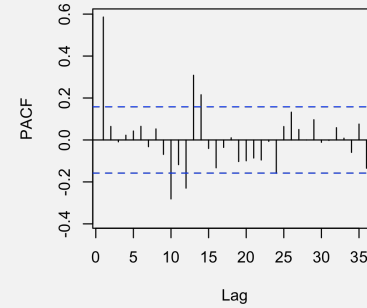
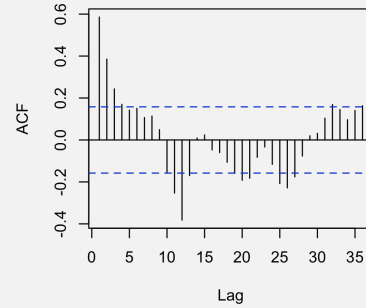
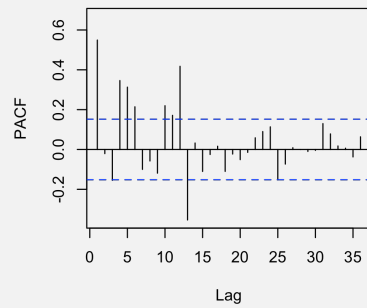
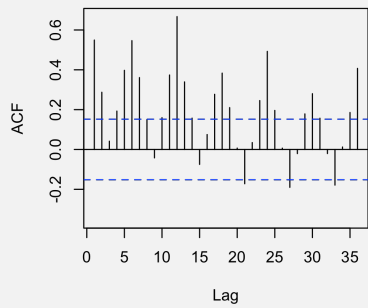
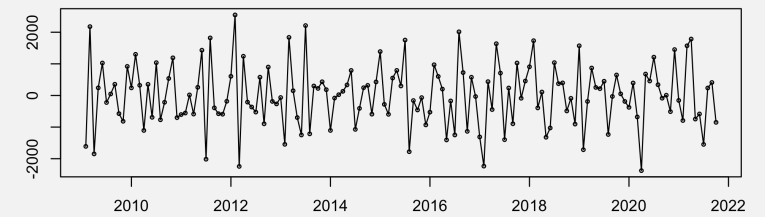
Diagnóstico da Série Temporal



Diagnóstico após diferenciação sazonal (lag = 12)



Diferenciação sazonal (lag = 12) e primeira diferenciação (lag = 1)



# ESTACIONARIEDADE

```
KPSS Unit Root Test
#####
##
Test is of type: mu with 4 lags.
##
Value of test-statistic is: 0.0199
##
Critical value for a significance level of:
10pct 5pct 2.5pct 1pct
critical values 0.347 0.463 0.574 0.739
```

```
Warning in adf.test(diff(diff(ts_train, lag = 12),
differences = 1)): p-value smaller than printed p-value
##
Augmented Dickey-Fuller Test
##
data: diff(diff(ts_train, lag = 12), differences = 1)
Dickey-Fuller = -6.9048, Lag order = 5, p-value = 0.01
alternative hypothesis: stationary
```

Ambos os testes confirmam a estacionariedade da série após a diferenciação de ordem 1 e 12

## PROPOSTAS DE MODELOS

---

SARIMA

---

GARCH

---

ETS

---

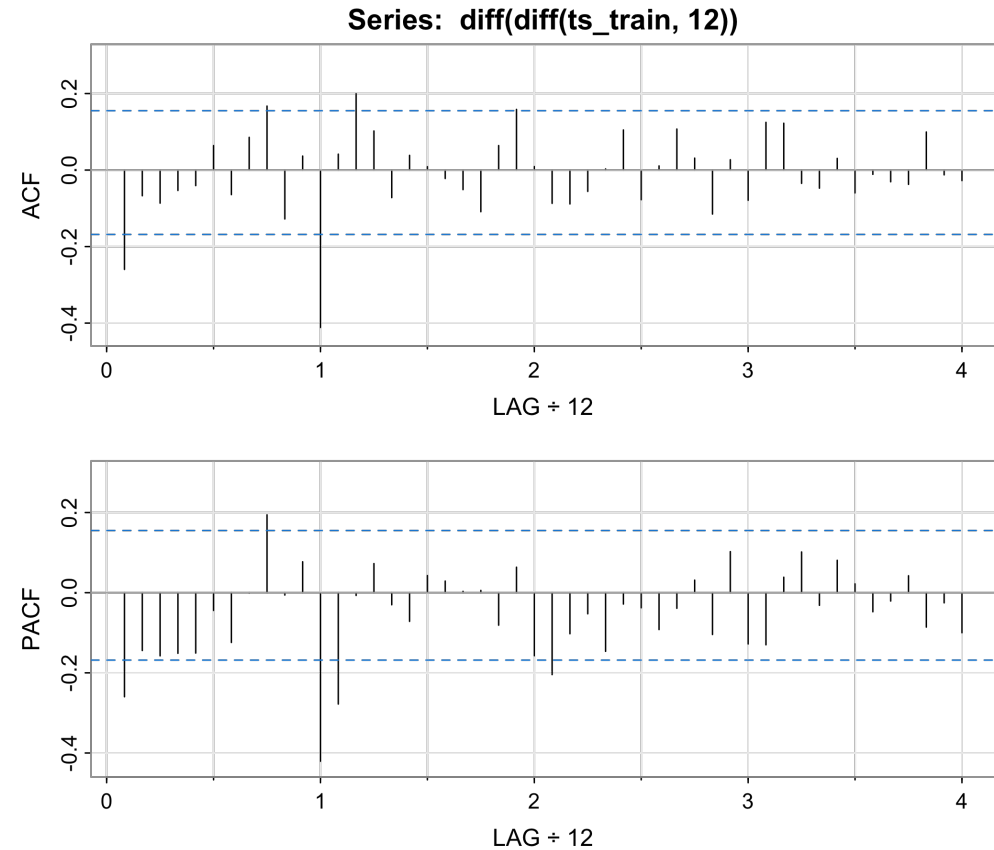
STLM



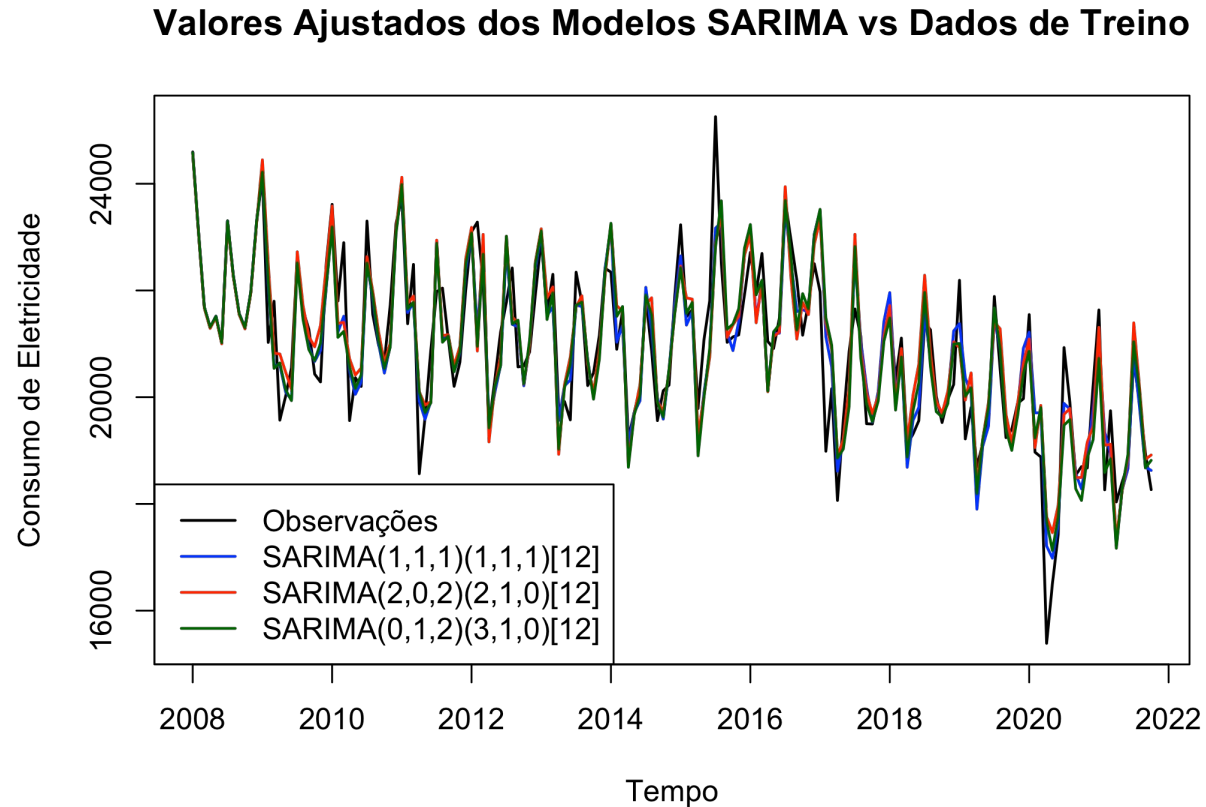
# SARIMA

3 modelos, tendo em conta:

- Análise ACF e PACF
  - SARIMA(1,1,1)(1,1,1)[12]
- Análise sistemática, critério AICc
  - SARIMA(2,0,2)(2,1,0)[12]
- Análise sistemática, critério AICc, com  $d = D = 1$ 
  - SARIMA(0,1,2)(3,1,0)[12]

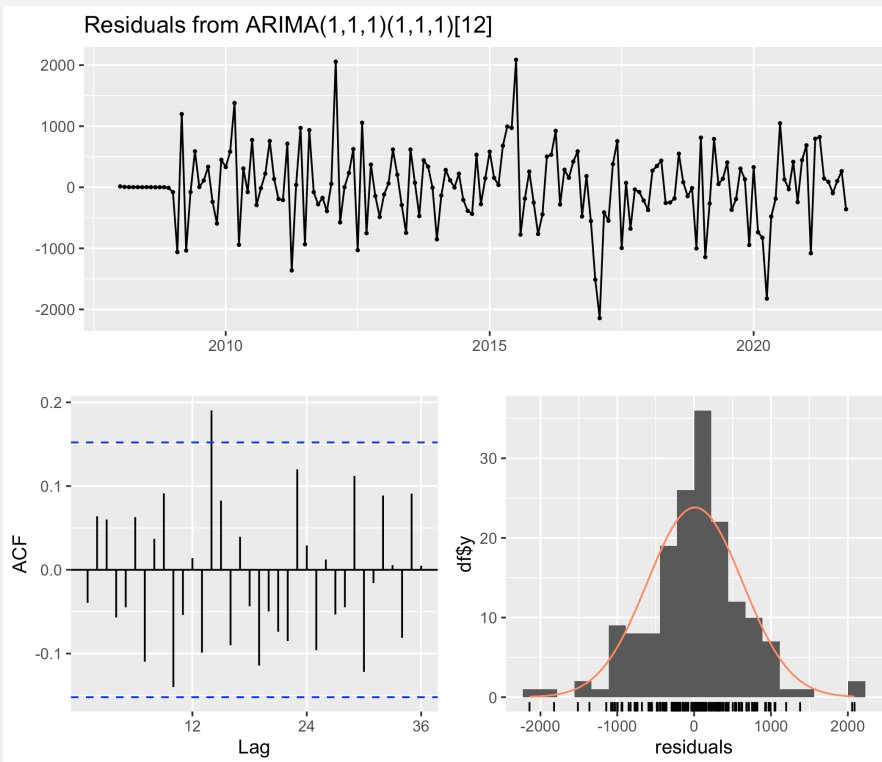


# SARIMA



- Apesar dos modelos parecerem ajustar-se bem aos dados de treino, todos os testes de Ljung-Box revelaram autocorrelação significativa nos resíduos
- Outras variantes foram testadas, mas apresentaram desempenhos ainda inferiores

# SARIMA



```
Ljung-Box test
```

```
##
```

```
data: Residuals from ARIMA(1,1,1)(1,1,1)[12]
```

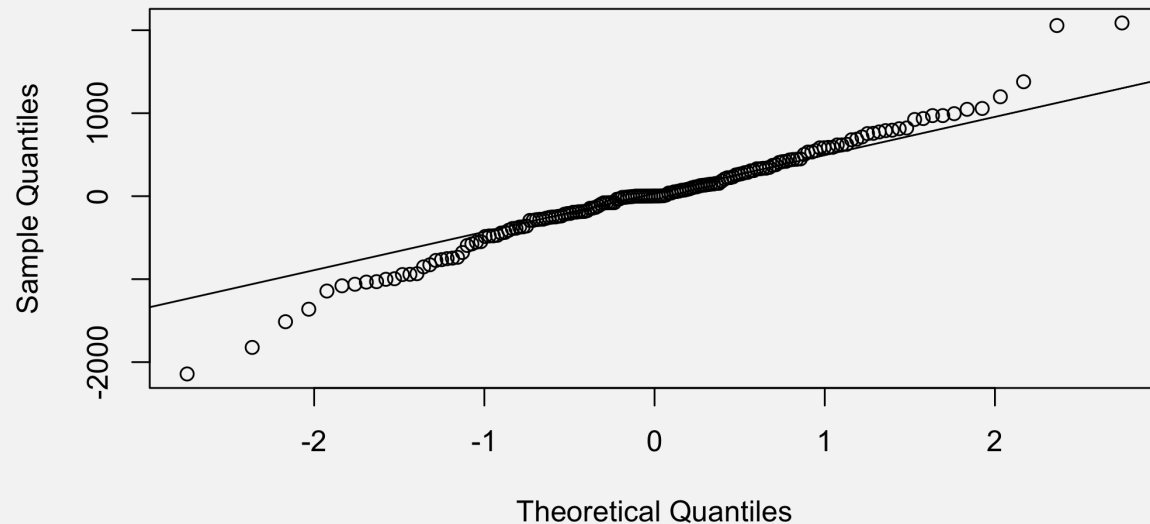
```
Q* = 31.304, df = 20, p-value = 0.0513
```

```
##
```

```
Model df: 4. Total lags used: 24
```

# SARIMA

Q-Q Plot dos Resíduos do SARIMA(1,1,1)(1,1,1)[12]



```
Shapiro-Wilk normality test
##
data: residuals(fit_arima)
W = 0.97568, p-value = 0.005097

Asymptotic one-sample Kolmogorov-Smirnov test
##
data: residuals(fit_arima)
D = 0.083543, p-value = 0.1969
alternative hypothesis: two-sided
```

# SARIMA

```
Series: ts_train
ARIMA(1,1,1)(1,1,1)[12]
##
Coefficients:
ar1 ma1 sar1 sma1
0.4499 -0.8689 0.0579 -0.9983
s.e. 0.1252 0.0830 0.0894 0.2415
##
sigma^2 = 424234: log likelihood = -1221.8
AIC=2453.6 AICc=2454.01 BIC=2468.75

ar1 ma1 sar1 sma1
ar1 1.000000000 -0.81456826 0.14406044 -0.005351725
ma1 -0.814568259 1.000000000 -0.14824210 0.004377140
sar1 0.144060443 -0.14824210 1.000000000 -0.022012213
sma1 -0.005351725 0.00437714 -0.02201221 1.000000000
```

Apesar de:

- Insignificância de *sar1*
- Correlação *ma1* e *ar1*

Foi mantida a estrutura completa do modelo, dado ser o compromisso mais aceitável entre qualidade do ajustamento, simplicidade estrutural e cumprimento dos pressupostos essenciais

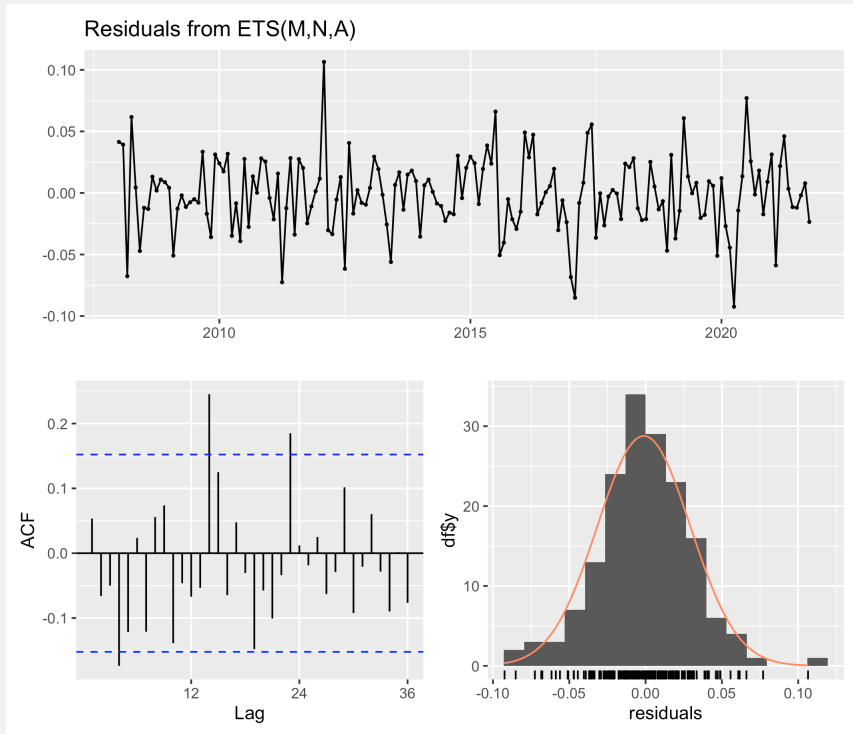


# GARCH

- Ajustado o modelo SARIMA, foi verificada a presença de efeitos ARCH nos resíduos do modelo

```
ARCH LM-test; Null hypothesis: no ARCH effects
##
data: residuals(fit_arima)
Chi-squared = 9.7974, df = 12, p-value = 0.6337
```

- Não faz sentido ajustar um modelo GARCH



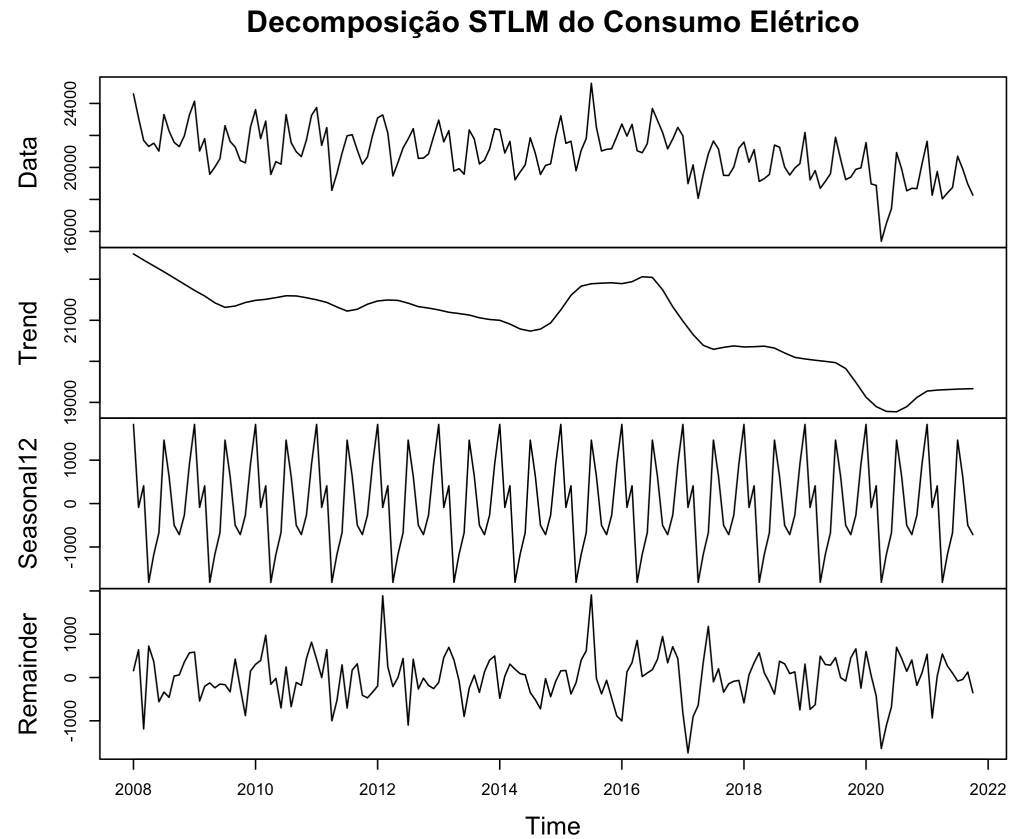
```
Ljung-Box test
##
data: Residuals from ETS(M,N,A)
Q* = 47.728, df = 24, p-value = 0.002727
##
Model df: 0. Total lags used: 24
```

# ETS

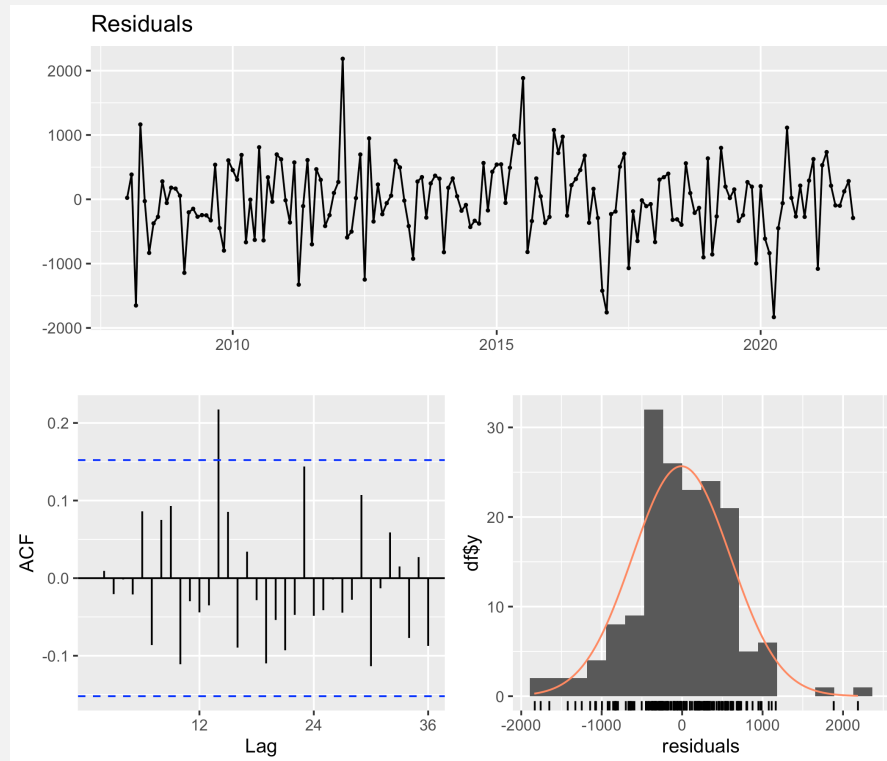
- Ajuste do modelo com o critério AICc
  - ETS(M,N,A)
- A elevada autocorrelação dos resíduos levou à rejeição de modelos desta família como adequados para representar a série

# STLM

- Decomposição STL com ARIMA ajustado aos resíduos dessazonalizados
  - STLM(ARIMA)



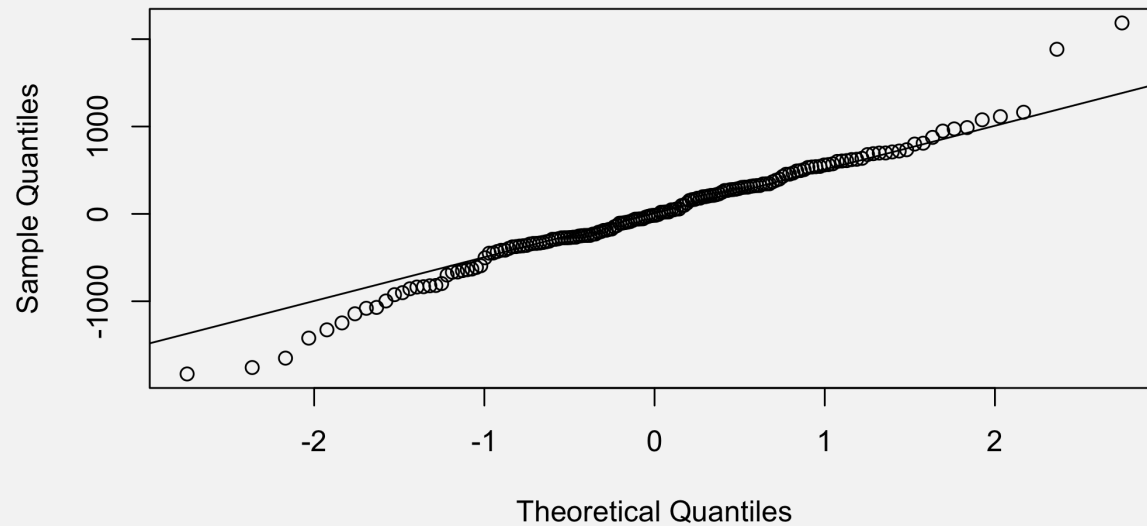
# STLM



```
Ljung-Box test
##
data: Residuals
Q* = 29.546, df = 24, p-value = 0.2003
##
Model df: 0. Total lags used: 24
```

# STLM

Q-Q Plot dos Resíduos do STLM(ARIMA)

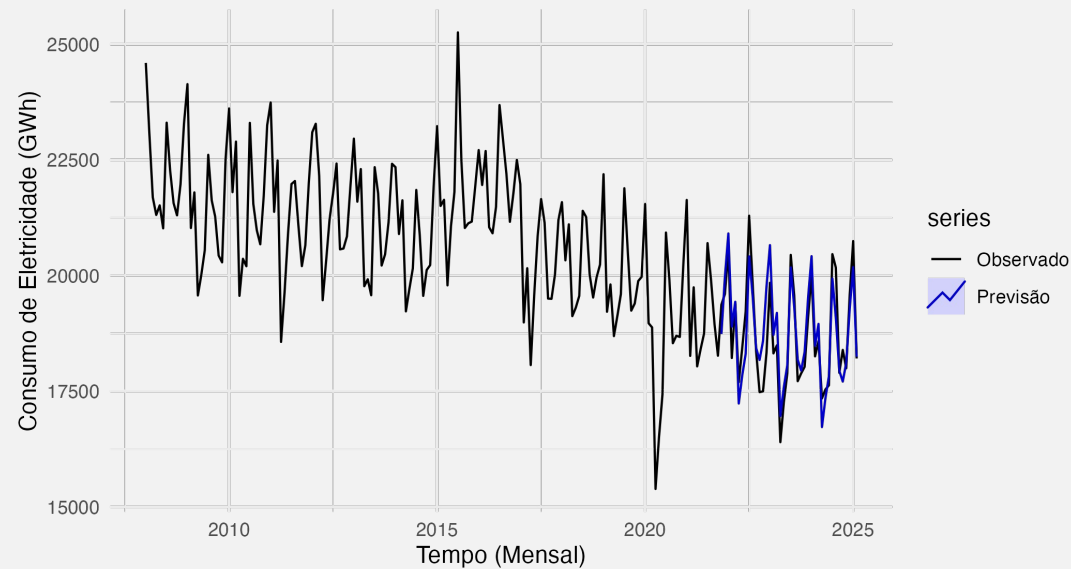


```
Shapiro-Wilk normality test
##
data: residuals(fit_stlm)
W = 0.9794, p-value = 0.01422

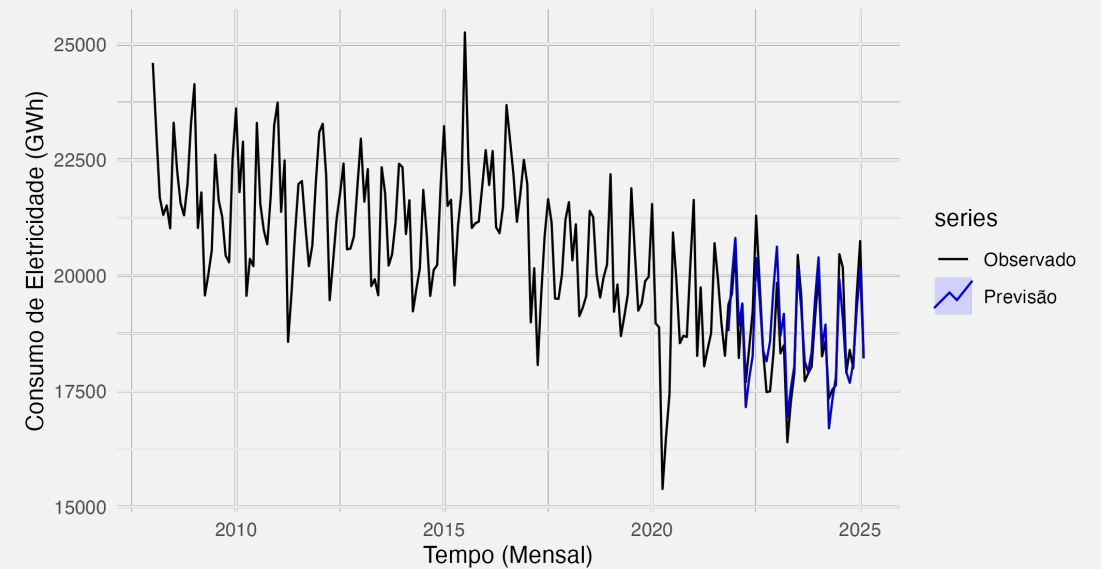
Asymptotic one-sample Kolmogorov-Smirnov test
##
data: residuals(fit_stlm)
D = 0.073808, p-value = 0.3263
alternative hypothesis: two-sided
```

# PREVISÃO DE OBSERVAÇÕES FUTURAS

Previsão vs Série Completa - SARIMA(1,1,1)(1,1,1)[12]



Previsão vs Série Completa - STLM(ARIMA)



## PREVISÃO DE OBSERVAÇÕES FUTURAS

|                 | ME        | RMSE     | MAE      | MPE      | MAPE    | MASE    | ACF1     | Theil's U |
|-----------------|-----------|----------|----------|----------|---------|---------|----------|-----------|
| SARIMA (Treino) | 10.45888  | 617.0807 | 448.1913 | −0.02917 | 2.16571 | 0.54375 | −0.03951 | NA        |
| SARIMA (Teste)  | −52.21608 | 562.8728 | 465.7346 | −0.35018 | 2.48337 | 0.56503 | 0.58261  | 0.42074   |

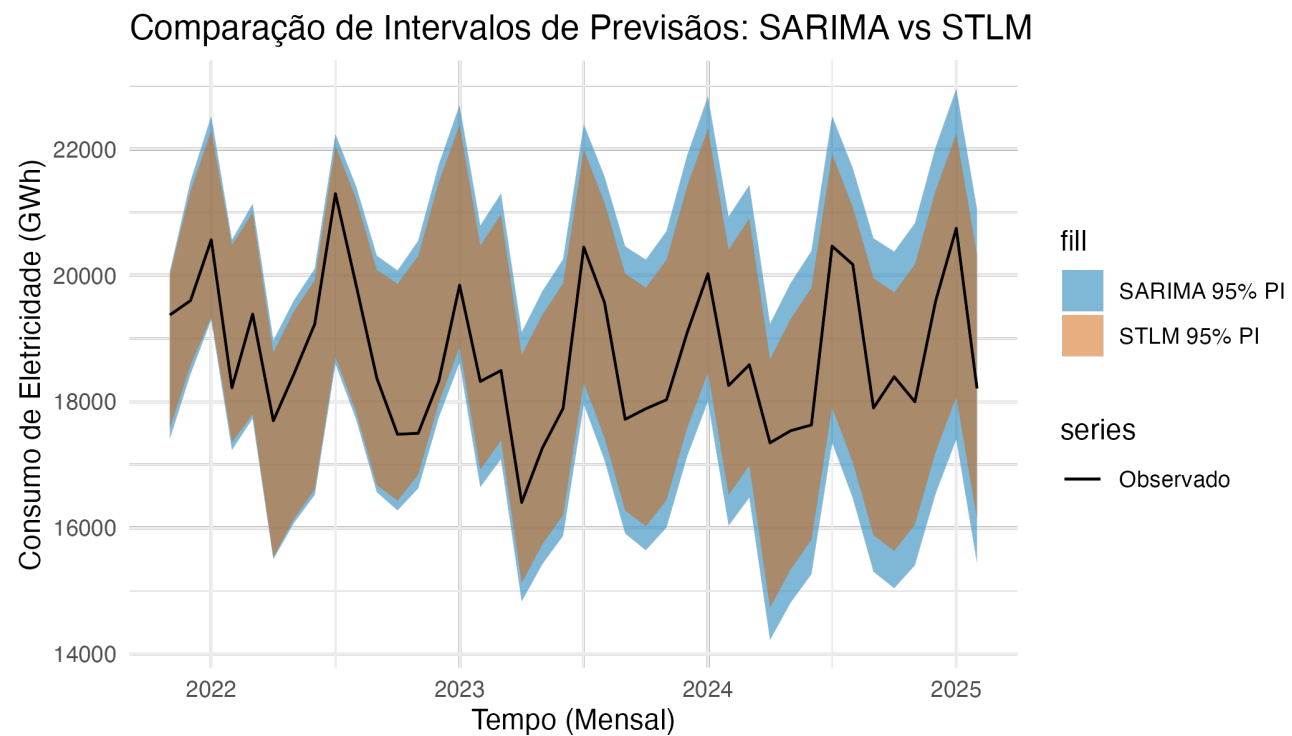
  

|               | ME        | RMSE     | MAE      | MPE      | MAPE    | MASE    | ACF1     | Theil's U |
|---------------|-----------|----------|----------|----------|---------|---------|----------|-----------|
| STLM (Treino) | −8.74319  | 607.9242 | 464.0256 | −0.12314 | 2.23067 | 0.56296 | −0.00060 | NA        |
| STLM (Teste)  | −26.13118 | 563.2452 | 460.1395 | −0.21075 | 2.45174 | 0.55824 | 0.58367  | 0.42183   |



## PREVISÃO DE OBSERVAÇÕES FUTURAS

- Ambos os modelos com cobertura total
- Amplitude média:
  - SARIMA: 4378
  - STLM: 3639



# DISCUSSÃO

## **Modelo escolhido:** STLM(ARIMA)

- Ambos os modelos apresentam desempenho preditivo equivalente
- Apresenta intervalos de previsão mais estreitos com igual nível de cobertura
- Resíduos com distribuição mais próxima da normalidade
- Maior evidência de ausência de autocorrelação nos resíduos