**Previsão de Séries Temporais do Consumo de Energia Elétrica**

**Da Região Sudeste do Brasil**

Guilherme Forindo Afonso¹\*;Felipe Pinto da Silva2

1 Aluno formado em Engenharia Elétrica pela Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP de Ilha Solteira - SP

2 Aluno de Doutorado do programa de pós-graduação em Economia pela Universidade Estadual de Campinas - UNICAMP. Mestre em Ciências Econômicas. Rua Pitágoras, 353 - Cidade Universitária, Campinas - SP, 13083-857. [dasilva.felipe@outlook.com](mailto:dasilva.felipe@outlook.com)

\*autor correspondente: [guilhermeflorindoafonso@gmail.com](mailto:guilhermeflorindoafonso@gmail.com)

|  |  |  |
| --- | --- | --- |
| **Aluno**: Guilherme Florindo Afonso | | **Data início curso**: 19/04/2021 |
| **Orientador**: Felipe Pinto | | **Defesa em:**  12/2022 |
| **Curso**: MBA Data Science e Analytics | **Modalidade**: Distância | Turma: DSA211 |

**Previsão de séries temporais do consumo de energia elétrica da região Sudeste do Brasil.**

1. **Resumo**

Tendo o conhecimento de que o consumo de energia elétrica no país todo vem crescendo de uma maneira considerável a cada ano que passa, e como o Brasil possui uma indústria de energia como característica principal a livre concorrência, assim como a privatização desse tipo de serviço aos consumidores é algo notório, isso fez com que houvesse uma reestruturação dessa indústria afim de que se forneça um serviço de qualidade aos consumidores. Sendo assim, o objetivo desse trabalho foi utilizar a base de dados de consumo de energia elétrica da região Sudeste Brasileira (de 2004 até 2022) e em cima dela aplicar alguns dos principais modelos de Séries Temporais a fim de se fazer uma previsão de como será a linha gráfica desse consumo para daqui alguns anos à frente. Depois de ter feito todos os tipos de modelagem e com os resultados em mãos, pode-se concluir que alguns dos modelos utilizados tiveram um ótimo nível de acurácia, com resultados estatisticamente significativos e com baixo valor de erro médio percentual absoluto (*MAPE*). Entre os melhores modelos estão: ARIMA e o de Suavização Exponencial (ETS). Em contrapartida o modelo de Regressão Linear falhou bastante ao tentar capturar os valores ajustados para essa série temporal em questão.

1. **Introdução**

O consumo médio de energia elétrica de todo o país vem crescendo consideravelmente a cada ano que passa, e isso torna um dos principais problemas a serem discutidos no país principalmente no âmbito econômico. A privatização e a desregulamentação do sistema elétrico são os maiores responsáveis pela mudança de cenário desse sistema (Oliveira, 2017).

A principal característica da indústria de energia e do Brasil é a livre concorrência na compra e venda de energia elétrica, sendo isso um ponto indispensável para a desverticalização, isto é, atividades de geração, transmissão e geração de energia (GTD) desempenhadas de forma autônoma e independente (Campos, 2008). Como consequência disso, as últimas duas décadas foram marcadas como uma crescente evolução do setor de energia elétrica para que se tornasse uma forte estrutura de mercado (Oliveira, 2017).

Sendo assim, foi necessária uma enorme reestruturação desse sistema a fim de que houvesse uma indústria de energia elétrica de qualidade, desse modo, fornecendo uma energia de qualidade e segura para os consumidores. Uma das principais ferramentas utilizadas para atender esse problema atualmente são modelos de previsão de carga, muito usados tanto para previsões de curto-prazo quanto de longo-prazo.

Deve-se prestar muita atenção às técnicas de previsão de carga uma vez que os custos dos erros de previsão acabam por gerar um custo financeiro muito elevado, seja por excesso (*overforecast*) quanto por erro de previsão por falta (*underforecast*) (Oliveira, 2004). Com o intuito de se alcançar uma eficiência no sistema e uma economia nos custos de eficiência do sistema, diversas técnicas estão sendo aplicadas.

Este trabalho apresentará algumas técnicas de modelagem matemática usadas para a previsão do consumo elétrico da região Sudeste brasileira. Alguns modelos de series temporais (dos simples aos clássicos e mais robustos) bem como de Redes Neurais serão utilizadas a fim de alcançar esse propósito.

Para a capacidade preditiva desses modelos serão realizados testes preliminares com uma base de dados do consumo de energia elétrica (em MWh) do ano de 2004 ao primeiro semestre do ano de 2022.

Este trabalho tem como objetivo abordar o problema de previsão de séries temporais para 5 anos à frente do consumo de energia elétrica da região Sudeste do Brasil por meio das seguintes metodologias de modelagem: modelos de séries temporais (ARIMA, Holt-Winters, ETS, Regressão Linear, Suavização Exponencial Simples) e por modelos de Redes Neurais Artificiais.

1. **Material e Métodos**

Este trabalho será conduzido primeiramente através de uma coleta de dados do consumo de energia do Brasil, todos registrados pela Empresa de Pesquisa Energética – EPE. Assim como mostra o título deste trabalho de conclusão de curso, serão realizadas algumas técnicas para filtragem e modelagem dos dados coletados somente para a região Sudeste do Brasil. Para isso, softwares como Excel e R Studio serão de grande importância para o desenvolvimento deste projeto.

Com a posse dos dados, previamente tratados e modelados, aplicar-se-á, o uso de algumas técnicas para uma série temporal do ano de 2004 até o primeiro semestre de 2022, a fim de fazer uma previsão desta série. Para isso, os modelos utilizados serão:

1. Modelo HoltWinters;
2. Modelo AutoRegressivo integrado de Média Móvel, ARIMA;
3. Modelo de Suavização Exponencial Simples (SES);
4. Modelo de Regressão Linear;
5. Modelo ETS (Error, Trend, Season);
6. Rede Neural Artificial, RNA.

Com os resultados em mãos, será feito um estudo das ferramentas estatísticas e matemáticas a fim de analisar a eficácia dos modelos, e ver qual deles funcionará melhor para o estudo do caso, ou seja, qualquer apresentará o melhor modelo preditivo para esta série temporal.

1. **Resultados Esperados**

O principal resultado esperado para este Trabalho de Conclusão de Curso é, através dos modelos aplicados aos dados coletados, escolher um entre todos, que tenha uma maior robustez preditiva à medida que as épocas das previsões aumentam. Em cima disso, fazer um estudo para futuros investimentos da indústria de energia elétrica na Região Sudeste do Brasil

1. **Testes e Resultados**

Como solução do modelo proposto para fazer as previsões de séries temporais do consumo de energia elétrica da região sudeste do Brasil, foi testado na base de dados de consumo de energia elétrica, do período de 2004 até a atualidade, 6 tipos de modelos de previsão de séries temporais utilizando software Rstudio.

Uma nota para o que vem a seguir, em relação aos textos que estão escritos na fonte “Simplified Arabic Fixed” são correspondentes as linhas de códigos que foram digitadas nos algoritmos do software R Studio.

Antes de tudo, fez-se uma plotagem da base de dados já transformada em uma

série temporal:

Gráfico, Gráfico de linhas

Descrição gerada automaticamente

Figura 1: Série Temporal do consumo de energia elétrica (MWh) do Sudeste brasileiro.

Fonte: Dados originais de pesquisa.

Para fazer as análises preliminares desta série temporal, foi utilizado o uso função “decompose()” pertencente ao pacote “stats” do Rstudio, no qual decompõem a série temporal entre os componentes de tendência, sazonais e componentes aleatórios, representados pela seguinte equação:

Onde:

Yt = valor da série no instante t;

Tt = componente de tendência para o instante t;

St = componente sazonal para o instante t;

Εt = componente aleatória para o instante t;

Interface gráfica do usuário, Gráfico

Descrição gerada automaticamente

Figura 2: Decomposição da série temporal com seus respectivos components (tendência, sazonalidade e aleatórios.

Fonte: Dados originais de pesquisa.

Pode-se analisar através da imagem de decomposição dos componentes da série temporal (Figura 2) que existem fortes traços de tendência e sazonalidade. Além disso, deve-se analisar os correlogramas das funções ACF (função de autocorrelação) e PACF (função de autocorrelação parcial) da diferenciação da série temporal através da função “ggtsdisplay()” pertencente ao pacote “forecast”. A diferenciação da série temporal é importante para que se analise-a em sua forma estacionária (média em torno do 0).

Interface gráfica do usuário

Descrição gerada automaticamente

Figura 3: Difenciação da série temporal com suas autocorrelações e autocorrelações parciais.

Fonte: Dados originais de pesquisa.

* 1. **Testando a acurácia dos modelos**

Antes de iniciar o teste dos resultados dos modelos aplicados à base de dados da série temporal, é importante destacar as ferramentas e equações mais importantes para avaliar a acurácia dos mesmos.

Um dos métodos de avaliação mais comuns para este tipo de problema são as funções de métricas de erros, existem várias entre elas, e a escolha depende do objetivo em que se deseja observar a acurácia destas métricas.

No trabalho em questão a medida escolhida foi o *MAPE* (Mean Absolute Percentage Error):

onde Yt: é o valor real da série no instante de tempo t e é o valor previsto pelo modelo utilizado para a previsão da série temporal no instante de tempo t.

* 1. **Modelo Holt-Winters**

A primeira técnica de previsão utilizada foi o modelo Holt-Winters pertencente à família de previsão de suavização exponencial, no qual é capaz de trabalhar com série temporais com ambos os componentes de tendência e sazonalidade. Antes de aplicar este método, fez-se a divisão da série temporal em uma parte de treino para aplicação do modelo e uma parte de teste (últimas 12 observações) para sua validação utilizando a função “ts\_split()” com as seguintes linhas de código:

consumo\_split <- ts\_split(ts\_consumo, 12)

treino\_HW <- consumo\_split$train

teste\_HW <-consumo\_split$test

autoplot(ts\_consumo) +

autolayer(treino\_HW, series = "Treino") +

autolayer(teste\_HW, series = "Teste") +

scale\_color\_viridis\_d() +

theme\_bw()

Gráfico, Gráfico de linhas

Descrição gerada automaticamente

Figura 4: Série temporal dividida em treino e teste.

Fonte: Dados originais de pesquisa.

Feito a divisão da série temporal, bem como a plotagem do gráfico das duas partes (treino e teste), foi aplicado o modelo Holt-Winters:

consumo\_HW <- HoltWinters(treino\_HW)

consumo\_HW

Tabela 1: Parâmetros do modelo Holt-Winters

Holt-Winters exponential smoothing with trend and additive seasonal component.

Call:

HoltWinters(x = treino\_HW)

Smoothing parameters:

alpha: 0.6607503

beta : 0

gamma: 0.7385392

Fonte: Dados Originais de pesquisa.

Aplicando a função “forecast() do pacote {forecast} e fazendo a análise da acurácia do modelo:

fc\_HW <- forecast(consumo\_HW, 12)

accuracy(fc\_HW, teste\_HW)

Tabela 2: Acurácia do modelo.

ME RMSE MAE MPE MAPE MASE ACF1

Training set: -30391.67 463649.1 357473.3 -0.1951976 1.886715 0.534331 0.09479866

Test set: -650102.07 886355.8 773329.1 -3.2054136 3.779057 1.155929 0.48409679

Fonte: Dados originais de pesquisa.

Como mostra na Tabela 2, para o modelo Holt-Winters aplicado, tem-se um Erro Médio Percentual Absoluto (MAPE) de 1,89% para o modelo de treino e um erro de 3,78% para o teste.

Para melhor análise e visualização, fez-se uma plotagem da série temporal original junto da série ajustada e prevista:

test\_forecast(actual = ts\_consumo,

forecast.obj = fc\_HW,

test = teste\_HW)

Gráfico

Descrição gerada automaticamente com confiança média

Figura 5: Dados reais vs Ajustados e Previsto.

Fonte: Dados originais de pesquisa.

Plotando os valores previstos e ajustados, tem-se um melhor entendimento da performance do modelo. Nota-se que o modelo Holt-Winters fez um bom trabalho em capturar tanto as componentes sazonais e de tendência. Porém, pode-se notar que o modelo não acompanha os valores reais em alguns pontos de pico.

De um modo alternativo, o modelo pode ser treinado de para que se otimize os parâmetros alpha, beta e gamma. O “grid search” é uma boa ferramenta para otimizar os ajustes desses parâmetros. O algoritmo é baseado em aplicar um conjunto de valores para cada parâmetro do modelo e fazer um teste com diferentes combinações de modo que se busque a minimização dos critérios de erro que será utilizada no modelo final.

Começando com uma busca rasa (shallow Search):

shallow\_grid <- ts\_grid(treino\_HW,

model = "HoltWinters",

periods = 6,

window\_space = 6,

window\_test = 12,

hyper\_params = list(alpha = seq(0,1,0.1),

beta = seq(0,1,0.1),

gamma = seq(0,1,0.1)),

parallel = TRUE,

n.cores = 8

)

shallow\_grid$grid\_df[1:10,]

plot\_grid(shallow\_grid)

Gráfico

Descrição gerada automaticamente

Figura 6: 100 melhores modelos dos parâmtros Holt-Winters.

Fonte: Dados originais de pesquisa.

O “plot\_grid” mostra uma visão intuitiva de um intervalo ideal de valores para cada parâmetro. Por padrão, a função está evidenciando os 100 melhores modelos em que os intervalos dos parâmetros são:

- alpha: 0,1 a 0,2

- beta: 0 a 0,2

- gamma: 0,3 a 0,9

Assim, pode ser aplicado os intervalos dos hiperparâmetros para um “grid search” profundo (deep grid):

deep\_grid <- ts\_grid(treino\_HW,

model = "HoltWinters",

periods = 6,

window\_space = 6,

window\_test = 12,

hyper\_params = list(alpha = seq(0.1,0.2,0.01),

beta = seq(0,0.2,0.01),

gamma = seq(0.3,0.9,0.01)),

parallel = TRUE,

n.cores = 8)

plot\_grid(deep\_grid)

Gráfico

Descrição gerada automaticamente

Figura 7: Hiperparêmtros ajustados pelo grid search.

Fonte: Dados originais de pesquisa,

O último passo deste processo é aplicar os valores ótimos dos parâmetros de suavização do “grid model” baseado nesta busca:

consumo\_HW\_grid <- HoltWinters(treino\_HW,

alpha = deep\_grid$alpha,

beta = deep\_grid$beta,

gamma = deep\_grid$gamma)

fc\_HW\_grid <- forecast(consumo\_HW\_grid, h= 12)

accuracy(fc\_HW\_grid, teste\_HW)

Tabela 2: Acurácia do modelo com parâmetros ajustados.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

ME RMSE MAE MPE MAPE MASE ACF1

Training set -121232.22 596485.3 444372.2 -0.6957118 2.376366 0.6642226 0.6433811

Test set -37473.02 533411.3 437872.8 -0.2059915 2.124061 0.6545076 0.4084597

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Fonte: Dados originais de pesquisa.

Comparando a acurácia do modelo inicial com o de ajustes de parâmetros através do “grid search”, pode-se notar que houve uma redução do erro absoluto percentual médio (MAPE) tanto do modelo de treino, quanto do modelo de teste, indicando uma melhoria significativa.

Gráfico

Descrição gerada automaticamente

Figura 8: Dados Reais vs Ajustados e Previstos com ajuste de parãmentros.

Fonte: Dados originais de pesquisa.

Visualmente, o modelo de previsão pode ser analisado em conjunto com os intervalo de confiança (80% e 95%):

Interface gráfica do usuário, Gráfico, Gráfico de linhas

Descrição gerada automaticamente

Figura 9: Valores observados e previstos.

Fonte: Dados originais de pesquisa.

* 1. **Modelo de Suavização Exponencial Simples (SES)**

A segunda técnica de previsão utilizada foi o modelo de Suavização Exponencia Simples (SES). É uma das principais classes de modelos usados em previsão de série temporais. Como o próprio nome diz, é um caso particular mais simples de suavização exponencial. Este tipo de modelo serve quando não existe nenhum padrão claro de sazonalidade ou tendência na série.

Através da análise preliminar da decomposição da série temporal em questão, sabe-se que existe tendência e sazonalidades evidentes, por esse mesmo motivo, este modelo será testado para que se comprove o nível de acurácia.

Do mesmo modo em que foi feito com o modelo Holt-Winters, dividiu-se a série em modelos para treino e para teste.

treino\_SES <- consumo\_split$train

teste\_SES <- consumo\_split$test

fc\_consumo\_SES <- ses(treino\_SES, h = 12, initial = "optimal")

fc\_consumo\_SES$model

Tabela 3: Parâmetros do Modelo de Suavização Exponencial Simples (SES)

## Simple exponential smoothing

##

## Call:

## ses(y = treino\_SES, h = 12, initial = "optimal")

##

## Smoothing parameters:

## alpha = 0.9449

##

## Initial states:

## l = 15157190.7731

##

## sigma: 538807.1

##

## AIC AICc BIC

## 6636.931 6637.048 6646.958

Fonte: Dados originais de pesquisa.

accuracy(fc\_consumo\_SES, teste\_SES)

Tabela 4: Acurácia do Modelo SES

## ME RMSE MAE MPE MAPE MASE ACF1

## Training set 24514.56 536222.9 411799.0 0.09833068 2.195192 0.615534 0.0007841877

## Test set 524518.19 830833.4 689959.9 2.46026364 3.303300 1.031313 0.5570409470

Fonte: Dados originais de pesquisa.

test\_forecast(actual = ts\_consumo,

forecast.obj = fc\_consumo\_SES,

test = teste\_SES)

Gráfico

Descrição gerada automaticamente

Figura 10:Dados Reais vs Ajustados e Previstos (SES).

Fonte: Dados originais de pesquisa.

Pode-se analisar através da plotagem da previsão, a função “ses()” utiliza o treino para identificar o nível da série estimando o parâmetro alpha e o valor inicial do modelo. O nível do valor previsto é bem próximo do valor da última observação da série, uma vez que, o valor de alpha é próximo a 1 (alpha – 0,9449). Como objetivo do modelo de Suavização Exponencial Simples é prever o nível da série, o modelo não captura as oscilações a curto-prazo de tendência e sazonalidade.

Em caso de uma previsão plana, os intervalos de confiança fazem um papel importante, uma vez que o nível de incerteza é maior.

plot\_forecast(fc\_consumo\_SES) %>%

add\_lines(x = time(teste\_SES) + deltat(teste\_SES),

y = as.numeric(teste\_SES),

name = "Partição de teste") %>%

layout(title = "Consumo de energia elétrica (MWh)",

xaxis = list(range = c(2004, max(time(ts\_consumo)) +

deltat(ts\_consumo))),

yaxis = list(range = c(10000000,25000000)))

Gráfico, Gráfico de linhas

Descrição gerada automaticamente

Figura 11: Gráfico de previsão do modelo SES.

Fonte: Dados originais de pesquisa.

Como pode-se analisar pela plotagem de previsão, o teste está destro do intervalo de confiança de 80%.

* 1. **Modelo Auto-Regressivo Integrado de Média Móvel (ARIMA)**

O modelo de previsão ARIMA possui 3 componentes importantes: (p,q,d) que dizem respeito à ordem da parte auto-regressiva, ao grau de diferenciação e à ordem da média móvel respectivamente.

Assim, como os procedimentos anteriores, inicialmente divide-se a série temporal em parte de treino, e teste para a validação do modelo.

treino\_ARIMA <- consumo\_split$train

teste\_ARIMA <- consumo\_split$test

consumo\_ARIMA <- auto.arima(treino\_ARIMA,

d=1,

D=1,

stepwise = FALSE,

approximation = FALSE,

trace = TRUE)

consumo\_ARIMA

Tabela 5: Parâmetros do Modelo ARIMA.

## Series: treino\_ARIMA

## ARIMA(1,1,1)(0,1,1)[12]

##

## Coefficients:

## ar1 ma1 sma1

## 0.6264 -0.8670 -0.6762

## s.e. 0.1504 0.1096 0.0676

##

## sigma^2 estimated as 202300440499: log likelihood=-2831.79

## AIC=5671.57 AICc=5671.78 BIC=5684.68

Fonte: Dados originais de pesquisa.

A função “auto.arima()” retorna o modelo ARIMA que minimiza os valores AIC (AIC = 5671,57). Neste caso em questão, o melhor modelo proposto é o SARIMA(1,1,0)(0,1,1)[12]. E é válido ressaltar que esta função automaticamente calcula o número de diferenciações necessárias para que nossa série seja estacionaria.

Checando os resíduos:

checkresiduals(consumo\_ARIMA)

Tabela 6: Parâmetros residuais do teste Ljung-Box

Ljung-Box test

data: Residuals from ARIMA(1,1,1)(0,1,1)[12]

Q\* = 27.781, df = 21, p-value = 0.1465

Model df: 3. Total lags used: 24

Fonte: Dados originais de pesquisa.

O teste Ljung-Box mostra um p-value de 0,1456 > 0,05, portando, aceita-se a hipótese nula, na qual diz que os resíduos não são correlacionados a um nível de significância de 5%, ou seja, o modelo não exibe falha de ajuste.

fc\_consumo\_ARIMA <- forecast(consumo\_ARIMA, h = 12)

accuracy(fc\_consumo\_ARIMA, teste\_ARIMA)

Tabela 7: Acurácia do Modelo ARIMA

## ME RMSE MAE MPE MAPE MASE ACF1

Training set -12334.37 432219.1 313213.1 -0.09847485 1.648819 0.4681734 -0.04227188

Test set -44875.17 537494.3 432005.6 -0.24263237 2.094508 0.6457377 0.39932749

Fonte: Dados originais de pesquisa.

test\_forecast(ts\_consumo,

forecast.obj = fc\_consumo\_ARIMA,

test = teste\_ARIMA)

Gráfico

Descrição gerada automaticamente com confiança média

Figura 12: Valores reais vs Ajustado e Previstos (ARIMA).

Fonte: Dados originais de pesquisa.

**5.5 Modelo de Regressão Linear**

Para a aplicação do Modelo De Regressão linear, utilizou-se a função “tslm()” do pacote {forecast}, na qual faz uma transformação do objeto de série temporal em um modelo de regressão linear. Sendo assim, pode-se aplicar o componente de regressão junto dos outros recursos. Sendo assim, fez-se uma previsão das últimas 12 observações da série temporal usando tendência, o quadrado da tendência e a componente sazonal.

treino\_tslm <- consumo\_split$train

teste\_tslm <- consumo\_split$test

consumo\_tslm <- tslm(treino\_tslm ~ season + trend + I(trend^2))

fc\_tslm <- forecast(consumo\_tslm, h =12)

test\_forecast(ts\_consumo,

forecast.obj = fc\_tslm,

test = teste\_tslm)

accuracy(fc\_tslm, teste\_tslm)

Tabela 8: Acurácia do modelo de regressão linear.

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

MAPE

Training set 2.679655

Test set 5.366021

Fonte: Dado originais de pequisa.

Gráfico

Descrição gerada automaticamente com confiança baixa

Figura 13: Dados reais vs Ajustados e Previsto (regressão linear)

Fonte: Dados originais da pesquisa.

Através da análise visual gráfica, percebe-se que o modelo falhou bastante nos valores ajustados e na parte de previsão, em que nitidamente não conseguiu uma proximidade com os valores reais da série temporal original. Dentre os modelos testados, esse foi o menos preciso.

**5.6 Modelo ETS (Exponential Smoothing)**

O Modelo ETS como o próprio nome diz, faz parte dos modelos enquadrados dentre aqueles que usam a suavização exponencial. A diferença deste modelo é o fato dele aplicar 3 tipos de parâmetros:

- Erro (E): aditivo(A) ou multiplicativo(M);

- Tendência (T): nenhuma (N), aditiva (A), multiplicativa (M) ou amortecida (Ad ou Md).

- Sazonalidade (S): nenhum (N), aditiva (A) ou multiplicativa (M)

Aplicando o modelo e analisando sua acurácia:

treino\_ets <- consumo\_split$train

teste\_ets <- consumo\_split$test

consumo\_ets <- ets(treino\_ets)

summary(consumo\_ets)

Tabela 9: Parâmetros do modelo ETS (Exponencial Smoothing).

##ETS(M,Ad,M)

##

## Call:

## ets(y = treino\_ets)

##

## Smoothing parameters:

## alpha = 0.7845

## beta = 0.0001

## gamma = 0.0001

## phi = 0.9783

##

## Initial states:

## l = 14851528.9279

## b = 66866.4812

## s = 1.0044 1.0216 1.0153 1.0013 0.984 0.9642

## 0.9688 0.9854 1.0197 1.0208 1.0058 1.0087

##

## sigma: 0.0237

##

## AIC AICc BIC

## 6563.462 6567.062 6623.624

Fonte: Dados originais de pesquisa.

A função “ets()” nos mostra o tipo de série ETS que se ajusta melhor aos dados da série temporal original. Portanto, tem-se um modelo multiplicativo (M), de tendência aditiva amortecida com sazonalidade multiplicativa.

fc\_ets <- forecast(consumo\_ets, h =12)

accuracy(fc\_ets, teste\_ets)

test\_forecast(ts\_consumo,

forecast.obj = fc\_ets,

test = teste\_ets)

Tabela 10: Acurácia do modelo ETS.

ME RMSE MAE MPE MAPE MASE ACF1

Training set 15210.06 432536.4 322146.8 0.05053417 1.716937 0.4815270 -0.02819024

Test set 168533.23 459869.5 336682.9 0.76996497 1.609362 0.5032547 0.39002253

Fonte: Dados originais de pesquisa.

Gráfico

Descrição gerada automaticamente com confiança baixa

Figura 14: Valores reais vs Ajustados e Previstos (ETS).

Fonte: Dados originais de pesquisa.

**5.7 Rede Neural Artificial (RNA)**

Para a aplicação do modelo de previsão de séries temporais utilizando as Redes Neurais Artificiais, fez-se o uso da função “nnetar()”, no qual simula uma rede neural do tipo “Feed-Forward” com uma única camada escondida e os lags de entrada.

treino\_rn <- consumo\_split$train

teste\_rn <- consumo\_split$test

consumo\_rn <- nnetar(treino\_rn)

consumo\_rn

Tabela 11: Parâmetros do Modelo de Rede Neural Artificial (RNA)

Series: treino\_rn

Model: NNAR(2,1,2)[12]

Call: nnetar(y = treino\_rn)

Average of 20 networks, each of which is

a 3-2-1 network with 11 weights

options were - linear output units

sigma^2 estimated as 209273742978

Fonte: Dados originais de pesquisa.

fc\_rn <- forecast(consumo\_rn, h = 12)

autoplot(fc\_rn)

accuracy(fc\_rn, teste\_rn)

Fonte: Acurácia do Modelo RNA

ME RMSE MAE MPE MAPE MASE ACF1

Training set -151.0034 457139.0 342534.7 -0.06008024 1.818956 0.51200 0.0375377

Test set 447405.4586 581494.7 490841.7 2.15972800 2.371918 0.73368 0.2385419

Fonte: Dados originais de pesquisa.

test\_forecast(ts\_consumo,

forecast.obj = fc\_rn,

test = teste\_rn)

Uma imagem contendo Gráfico

Descrição gerada automaticamente

Figura 15: Dados reais vs Ajustados e Previstos.

Fonte: Dados originais de pesquisa.

**5.8 Comparando a acurácia dos modelos**

Tabela 12: Erro Médio Percentual Absoluto (%)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Holt-Winters** | **SES** | **ARIMA** | **Regressão Linear** | **ETS** | **RNA** |
| **Treino** | 2,38 | 2,19 | 1,65 | 2,68 | 1,72 | 1,82 |
| **Teste** | 2,12 | 3,30 | 2,1 | 5,37 | 1,61 | 2,37 |

Através da tabela acima, podemos concluir tranquilamente que os modelos que melhor se ajustaram à série temporal em questão foram os modelos ARIMA e ETS. Por outro lado, o modelo que apresentou o maior erro percentual absoluto médio foi o de Regressão Linear.

1. **Conclusões**

Tendo em vista de que qualquer projeto que for aplicado com base em um estudo feitos através de uma previsão de modelos, deve-se tomar total cuidado ao escolher o melhor modelo apresentado uma vez que, tratando-se de uma previsão do consumo de energia, existe uma margem de erro para esse tipo de aplicação. Deve-se prestar muita atenção às técnicas de previsão de carga uma vez que os custos dos erros de previsão acabam por gerar um custo financeiro muito elevado.

Após a coleta de dados e feito às devidas modelagens de previsão, segundo os valores de Erro Médio Percentual Absoluto da Tabela 12, temos como os modelos mais seguros para serrem utilizados nessa base de dados, os modelos ARIMA e ETS (Suavização Exponencial).

Em contrapartida, pode-se descartar sem dúvidas a utilização da Regressão Linear para esse em estudo em questão, uma vez que através da Figura 13, percebe-se que o modelo perdeu muita acurácia em relação aos valores reais e os valores previstos e ajustados, sendo assim, o pior entre os modelos que foram utilizados.

Sendo assim, é muito importante que se leve em consideração a parte estatística de cada modelo, observando atentamente sempre os parâmetros e as medidas de erros de previsão, evitando dessa maneira gastos e prejuízos financeiros investidos, e também pensando sempre nos consumidores.

1. **Referências Bibliográficas**

Fávero, L. P; Belfiore, P. 2017. Manual de Análise de Dados: Estatística e Modelagem Multivariada com Excel, SPSS e Stata. Editora GEN LTC, 1ª edição.

Morenttin, P. A. 2018. Análise de Séries Temporais: Modelos Lineares Multivariados (Volume 1). Editora: Blucher, 3ª edição.

Morenttin, P. A. 2020. Análise de Séries Temporais: Modelos Multivariados e Não Lineares (Volume 2). Editora: Blucher; 1ª edição.

Nielsen, A. 2021. Análise Prática de Séries Temporais: Predição com Estatística e Aprendizado de Máquina. Editora: Alta Books; 1ª edição.

Chollet, F. 2018. Deep Learning With R (English Edition). Editora: Manning; 1ª edição.

Goodfellow, I; Bengio. I; Courville. A. 2016. Deep Learning (Adaptive Computation and MAchine Learning Series) (English Edition). Editora: The MIT Press; 1ª edição.

Agarwall. C. C. 2018. Neural Networks and Deep Learning: A Textbook. Editora: Springer; 2018ª edição.

Bruce, P; Brune, Andrew; Gedeck, P. 2020. Pratical Statstics for Data Scientists: 50+ Essencial Concepts Using R and Python. Editora: O’Reilly; 2ª edição.

Empresa de Pesquisa Energética. 2022. Consumo Mensal de Energia Elétrica por Classe (regiões e subsistemas). <https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/consumo-de-energia-eletrica>