

MT_279 analysis

Margaret Reed

2023-05-07

```
library(tidyverse)
library(lubridate)
library(prophet)
library(MLmetrics)
library(scales)
```

```
source("make-model.R")
```

```
data <- get_account("MT_279")
```

```
data %>%
  ggplot(aes(x = ds, y = y)) +
  geom_line() +
  theme_minimal()
```



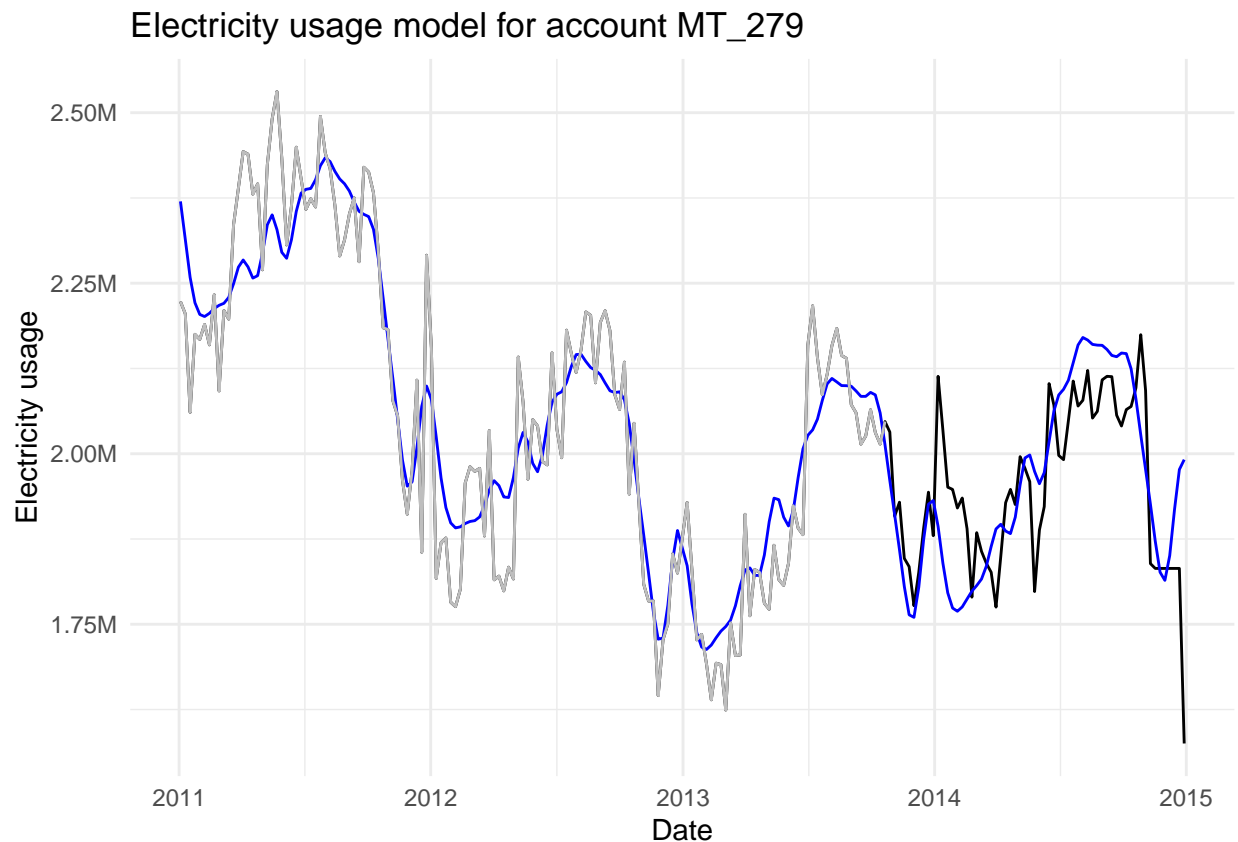
There appears to be an outlier so I looked at the end of the data

```
data <- data %>% slice(1: nrow(data) - 1)
```

```
train <- split_data(data)$train  
validation <- split_data(data)$validation  
test <- split_data(data)$test
```

```
m <- prophet(rbind(train, validation))  
future <- make_future_dataframe(m, periods = nrow(test), freq = "week")  
forecast <- predict(m, future)
```

```
full_join(data, forecast) %>%  
  ggplot() +  
  geom_line(aes(x = ymd(ds), y = y), color = "black") +  
  geom_line(aes(x = ymd(ds), y = yhat), color = "blue") +  
  theme_minimal() +  
  geom_line(data = data%>%slice_head(n=nrow(train)), aes(x = ymd(ds), y = y), color = "grey") +  
  scale_y_continuous(labels = label_number(scale = 0.000001, suffix = "M")) +  
  labs(  
    x = "Date",  
    y = "Electricity usage",  
    title = "Electricity usage model for account MT_279"  
  )
```



```
find_mape(forecast, test)$total
```

```
## [1] 0.04655262
```

```
find_mape(forecast, test)$first
```

```
## [1] 0.0358837
```

```
find_mape(forecast, test)$second
```

```
## [1] 0.03805001
```

```
find_mape(forecast, test)$third
```

```
## [1] 0.0689194
```

```
split_mape(forecast, test )%>%  
  pivot_longer(  
    cols = everything(),  
    names_to = "segment",  
    values_to = "error"  
  ) %>%  
  ggplot(aes(x = error, y = segment)) +  
  geom_boxplot() +  
  scale_x_continuous(label = label_percent()) +  
  labs(  
    x = "Error (MAPE)",  
    y = "Third of testing data",  
    title = "Boxplots of errors"  
  ) +  
  theme_minimal()
```

