TSA Competition

Aditi Jackson and Cara Kuuskvere

2024-03-28

Setup

First R code chunk is used for setting the options for all R code chunks. The choice echo=TRUE means both code and output will appear on report, include = FALSE neither code nor output is printed.

```
knitr::opts_chunk$set(
    message = FALSE,
    warning = FALSE,
    include = FALSE
)
library(lubridate)
```

```
##
## Attaching package: 'lubridate'
```

```
## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union
```

```
library(ggplot2)
library(forecast)
```

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

```
library(Kendall)
library(tseries)
library(outliers)
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr   1.1.4     v stringr 1.5.0
## v forcats 1.0.0     v tibble  3.2.1
## v purrr   1.0.2     v tidyr   1.3.0
## v readr   2.1.4
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(smooth)
```

```
## Loading required package: greybox
## Package "greybox", v2.0.0 loaded.
##
```

```
##
## Attaching package: 'greybox'
##
## The following object is masked from 'package:tidyr':
##
##     spread
##
## The following object is masked from 'package:lubridate':
##
##     hm
##
## This is package "smooth", v4.0.0
```

```
library(dplyr)
library(readr)
library(readxl)
library(zoo)
```

```
##
## Attaching package: 'zoo'
##
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

Data

Time Series

The initial plots show an weak exponential decay on the ACF and a strong correlation at lag 1 on the PACF. The TS plot shows at least two strong seasonal components; with seasonality within the years as well as seasonality within the months due to regular sinusoidal oscillations.

Decomposed data shows an increasing trend line and a strong seasonal component from the consistent wave-like pattern. Residuals do not exhibit any clear patterns.

Forecasting with Naive methods

Models 6, 7 and 8 all yielded the same results. We will need to incorporate exogenous variable if we will improve our model further.

NEURAL NETWORKS

We fit three Neural Network models: one with humidity as a regressor (NNETAR1), one with temperature as a regressor (NNETAR1), and one with both temperature and humidity as regressors (NNETAR3). Of these three, NNETAR3 produced the best forecast compared to the vanilla model. The forecast results improve by incorporating exogenous variables. We will need to try a model that can capture the seasonal variance in humidity in temperature if we are to improve our forecast.

TBATS

The TBATS model with temperature and humidity as exogenous regressors produced the best forecast of all the models attempted.