BAX 493A - Homework 1

—Asha Mairh, Kavish Hukmani, Keshore Suryanarayanan, Neon Zhang, Vinit Hasija

Deliverable 1

The exhaustive list of all columns in the dataset are stated in the below screenshot:

```
 [1] "yearweek"           "id_week"            "netsales_online"
 [4] "direct_mail"        "emails"             "phone_calls"
 [7] "online_budget"      "promotion"          "holiday"
[10] "avg_cloud_index"    "avg_temp"           "avg_wind_speed"
[13] "avg_rainfall_level" "avg_hours_sun"      "avg_snow_level"
```

Following are the selected columns for time series analysis, 1 per person:

- avg_cloud_index
- avg_temp
- avg_wind_speed
- avg_rainfall_level
- avg_hours_sun

For each variable, the time series is decomposed into trend, seasonality and a random component. While trend and seasonality are the useful signal components, the random component is noise.

The Holt-Winter method of exponential smoothing is used. Since there are 4 combinations of filters possible (assuming that level is ever-present), the 4 filters are experimented with each of the 5 variables in the R code. However, in this document, the plots and interpretations for only **Rainfall Level** are presented since it is depictive of the rest of the weather variables as well.

Deliverable 2

For each variable, we tried the following 4 models with all combinations.
beta → Trend coefficient
gamma → Seasonality coefficient

```
#1 - With trend and seasonality

 out1 = HoltWinters(data_ts, beta = TRUE, gamma = TRUE)
```

```
#2 - With trend and w/o seasonality
out2 = HoltWinters(data_ts, beta = TRUE, gamma = FALSE)

#3 - W/o trend and with seasonality
out3 = HoltWinters(data_ts, beta = FALSE, gamma = TRUE)

#4 - W/o trend and w/o seasonality
out4 = HoltWinters(data_ts, beta = FALSE, gamma = FALSE)
```

Deliverable 3

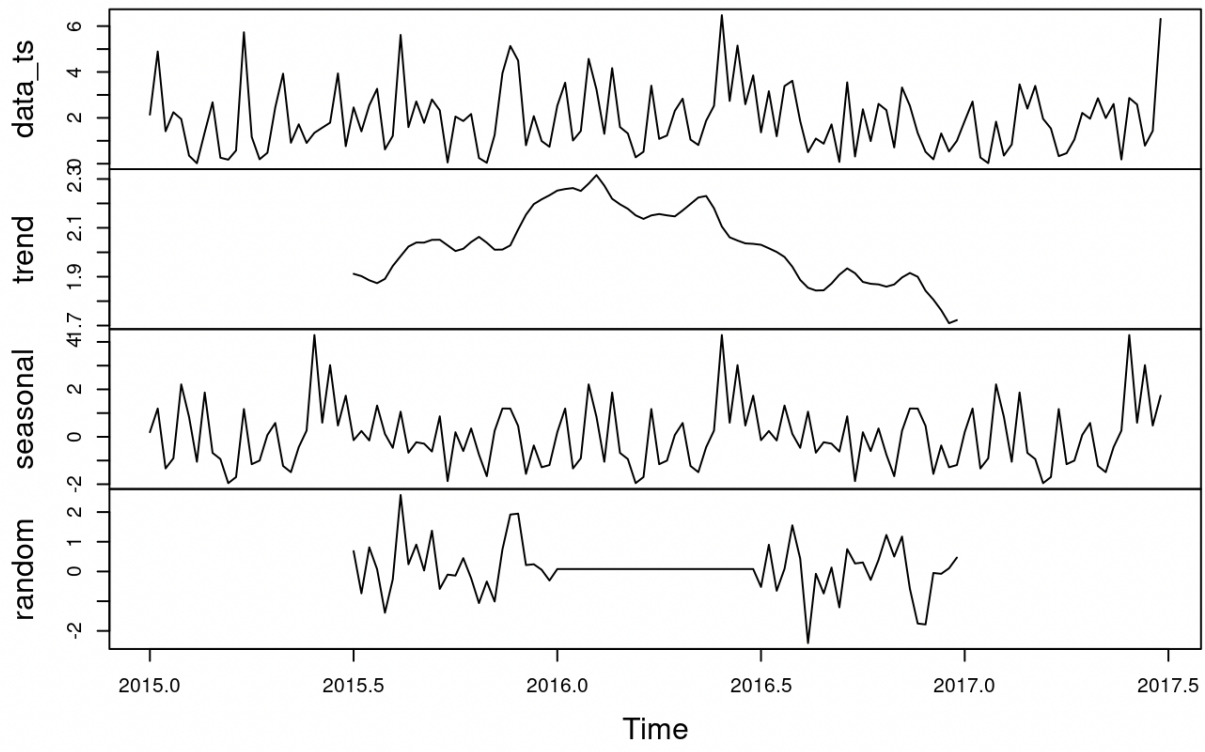

Fig 3.1 - Data Series, Trend, Seasonal plot for Rainfall Level

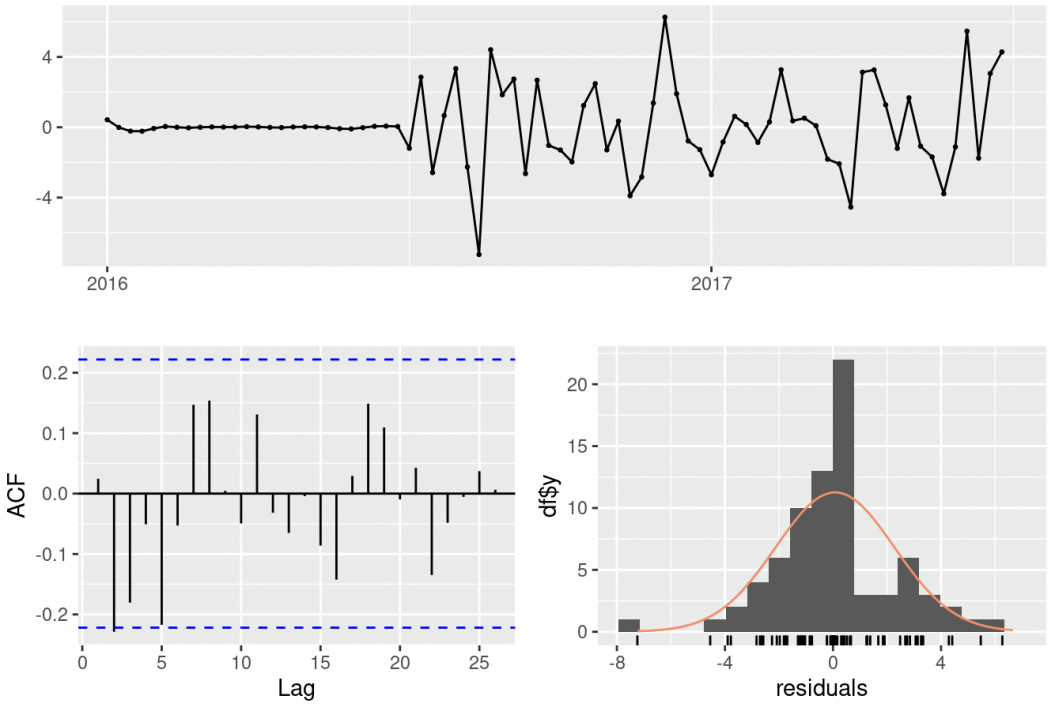

Fig 3.2 - With Trend & Seasonality

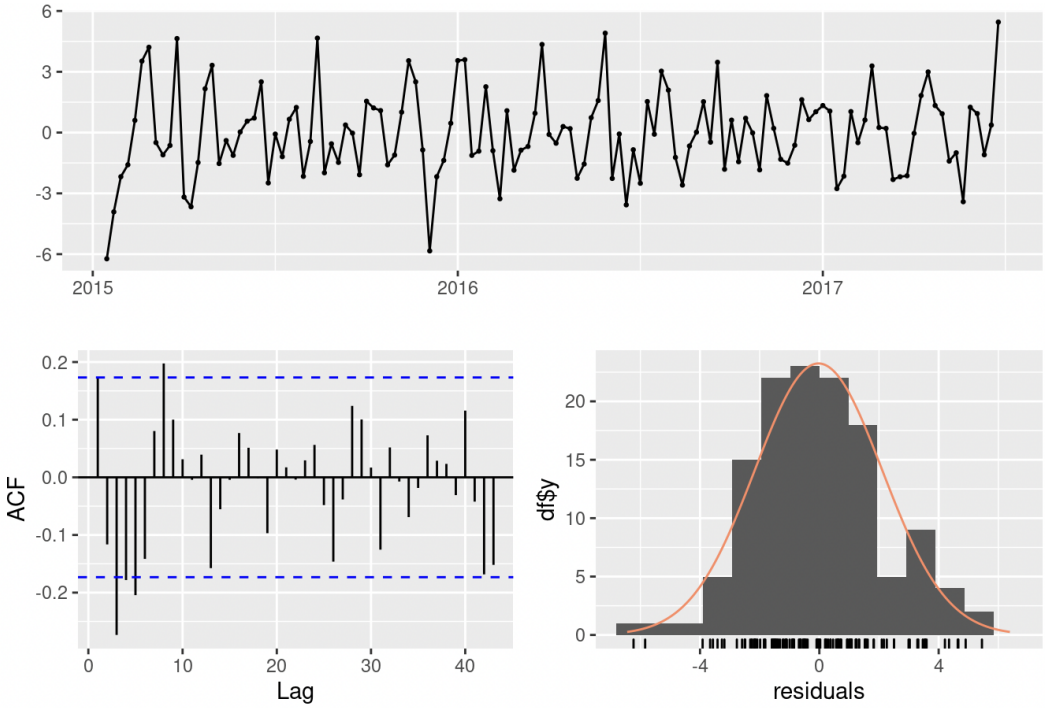

Fig 3.3 - With Trend & Without Seasonality

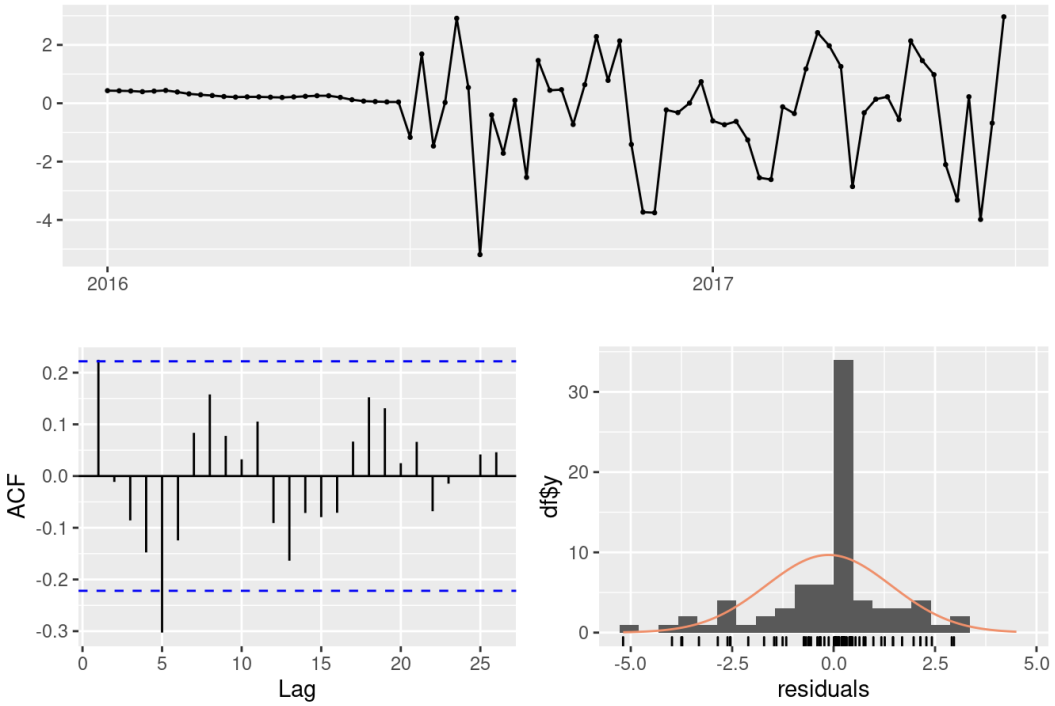

Fig 3.4 - Without Trend & With Seasonality

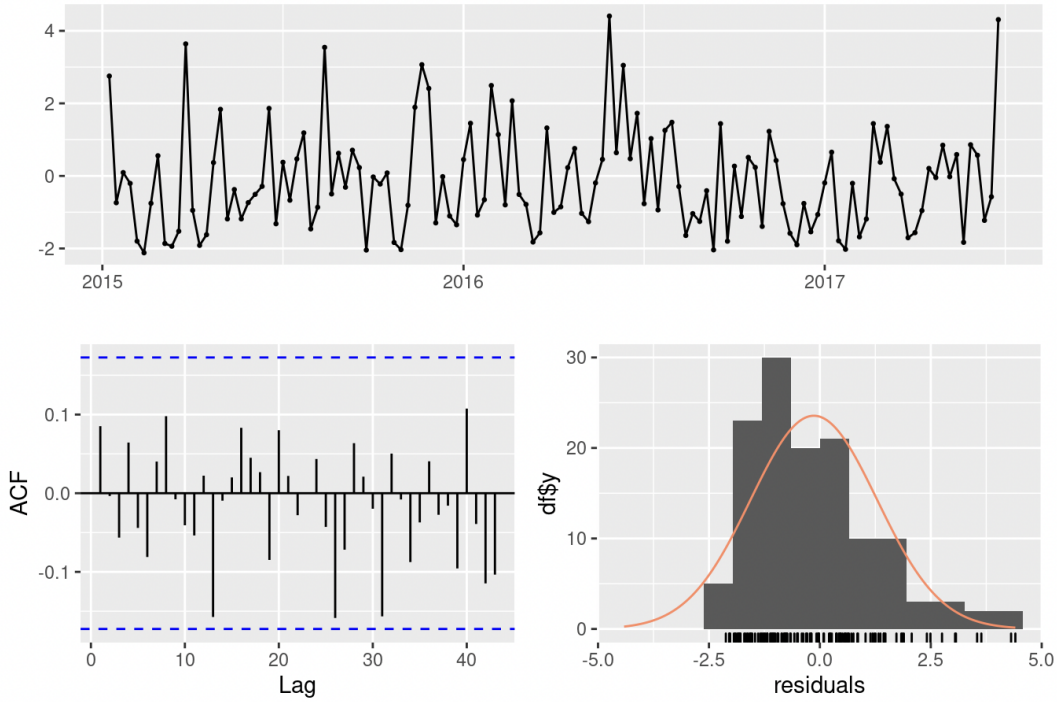

Fig 3.5 - Without Trend & Seasonality

Deliverable 4

Following are the out-of-sample forecasts for 26 weeks for the Rainfall level

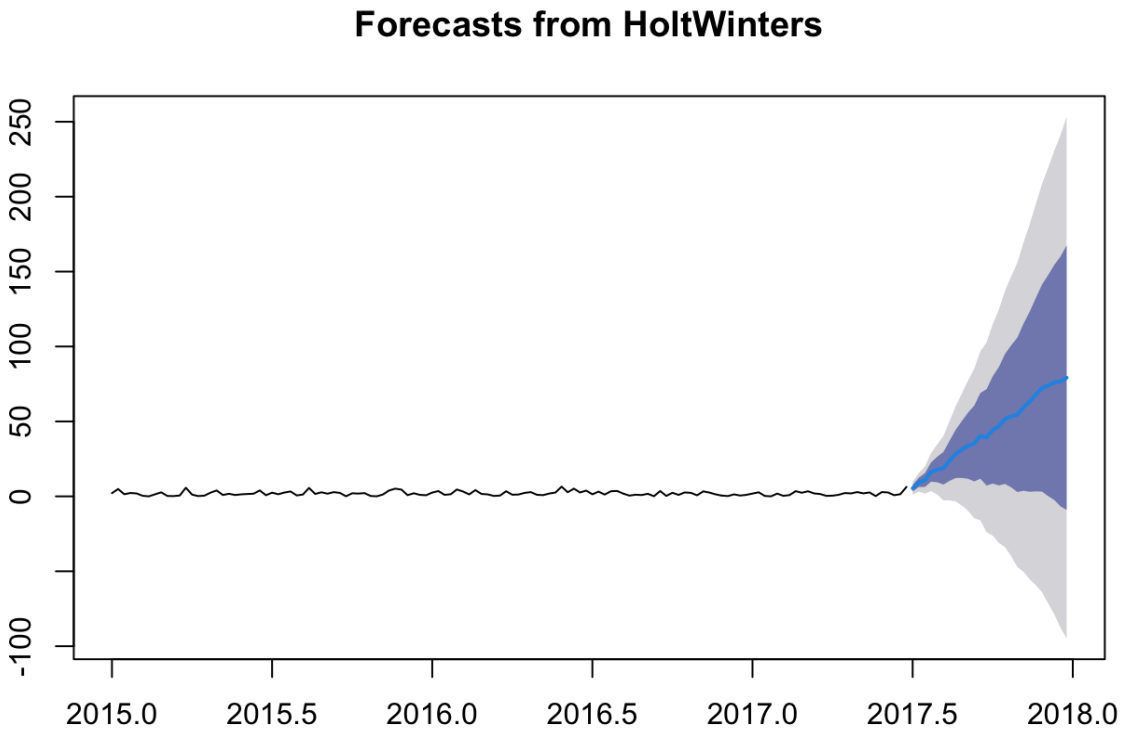

Fig. 4.1 - With Trend & Seasonality

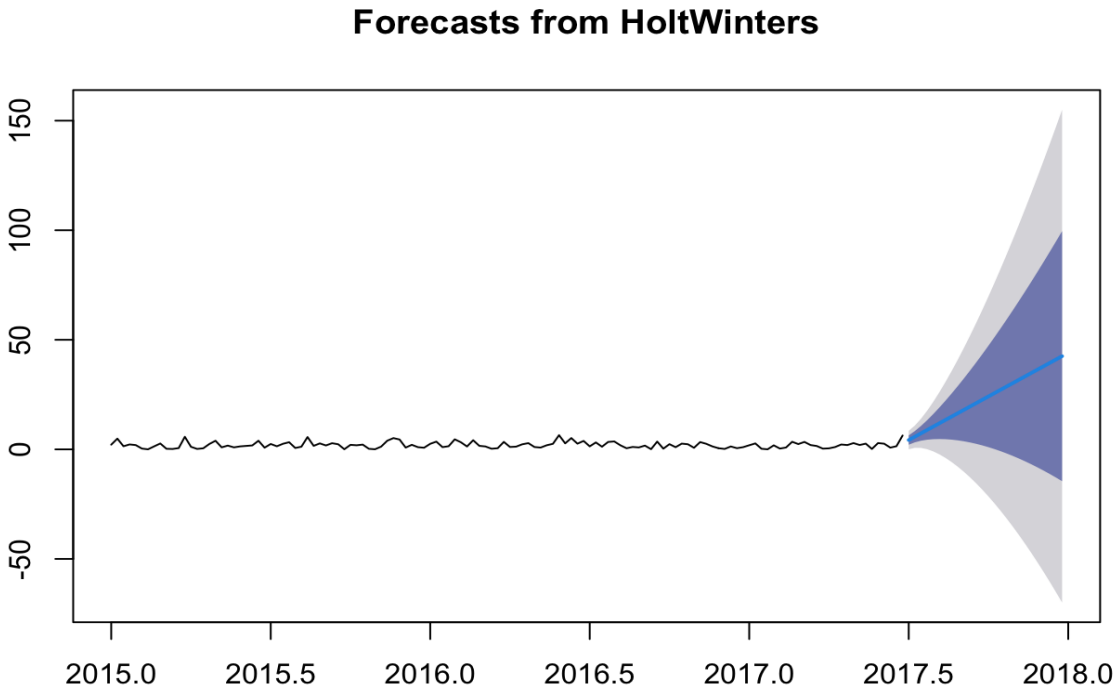

Fig. 4.2 - With Trend and w/o Seasonality

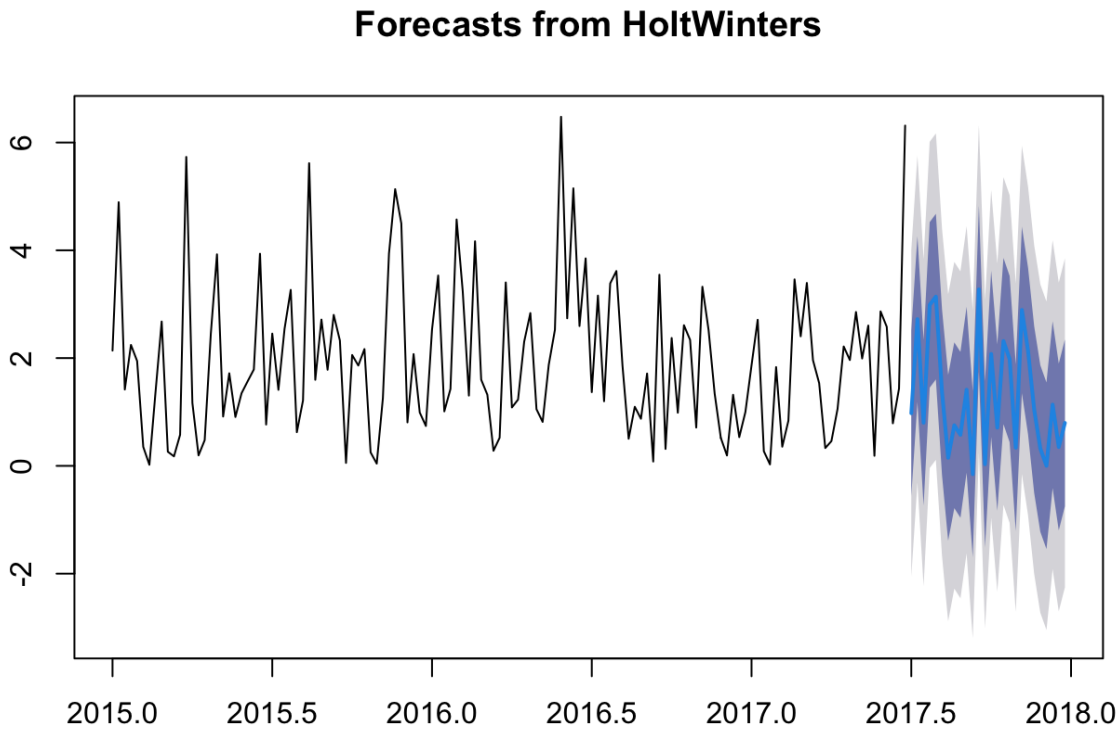

Fig. 4.3 - Without Trend & with Seasonality

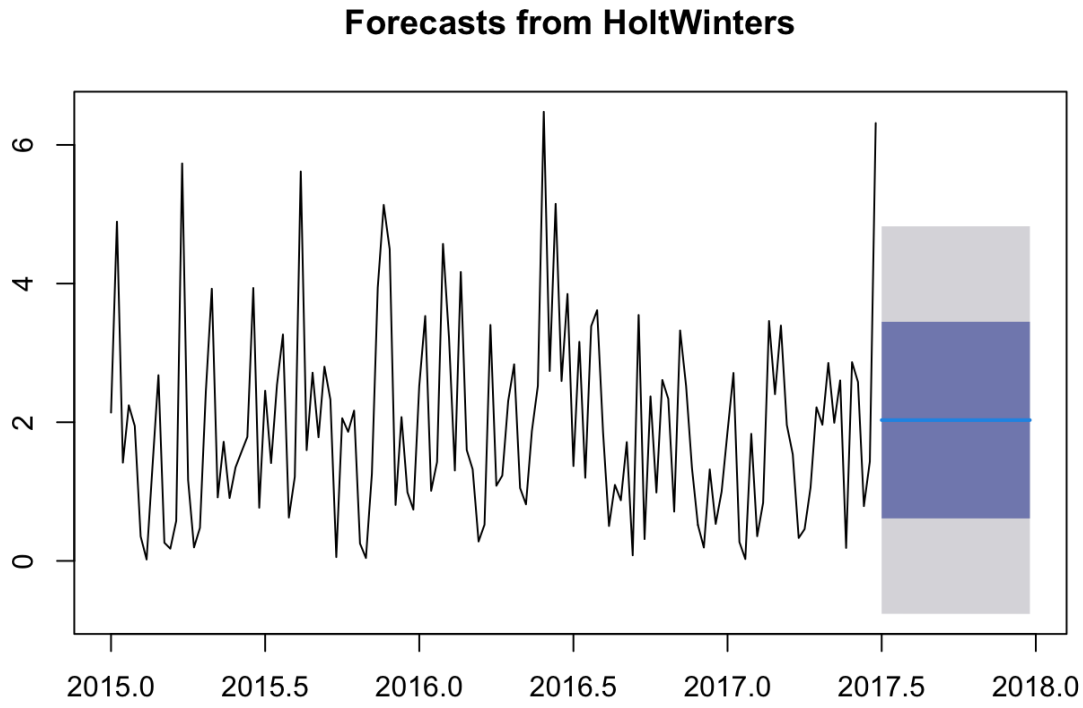

Fig. 4.4 - With Trend & Without Seasonality

Deliverable 5

The recommended forecasting model for all the weather variables considered in the analysis - cloud index, temperature, wind speed, rainfall level and sun hours - is the **Alpha-Gamma Filter** which accounts for level and seasonality, but not trend. There are two kinds of justifications for this - statistical and intuitive - as described below:

- **Statistical**: In the forecast plots, it is seen that the values predicted for the weather variables by all models other than the chosen one do not make sense. Take rainfall, for instance. In the original dataset, the values are between 0 and 6 but the Alpha-Beta-Gamma and Alpha-Beta models forecast values that touch 50 and 100, which are not realistic. Applying only the Alpha filter is also not helpful because there is seasonality present in the data which is not being accounted for in the predictions. Only the Alpha-Gamma model, which accounts for level and seasonality, makes sensible predictions and also follows a pattern that is consistent with the in-sample data.
- **Intuitive**: Weather variables are seasonal in nature. Yes, there could be trends due to the likes of global warming, but they are more observable across a longer period of time. For example, if the original dataset had, say, 10 years of data using which we are forecasting for 5 years, then there could be specific trends that the Holt-Winters model can capture and forecast. But the original dataset contains data for only 2 years using which we are making predictions for another 2 years, making it difficult to capture any trends. Which is why it makes sense to use an Alpha-Gamma filter that accounts for level and seasonality.