Junyu Yin G2101985L

1. Model Selection

First of all, let's show the plot of the original data to get some sense of it. From the following figure, we can see there are both distinct seasonal and trend components here. More specifically, we can observe that this time series has a seasonal period of 12 and increases year by year, which corresponds to that this data is for monthly anti-diabetic drug sales in Australia from 1992 to 2008.

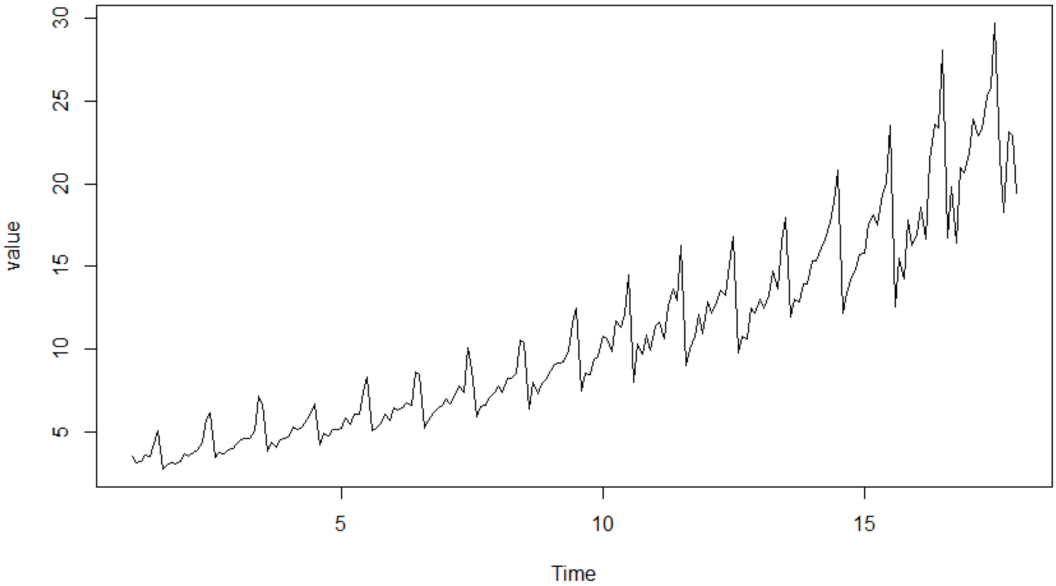

Figure 1. Time plot of the original time series.

Taking these factors into consideration, I decided to use the Holt-Winters' Trend and Seasonality (H-W) method to model this data. There are two main types of models, the additive and the multiplicative, that can be used here. The following figure shows the two fitted models.

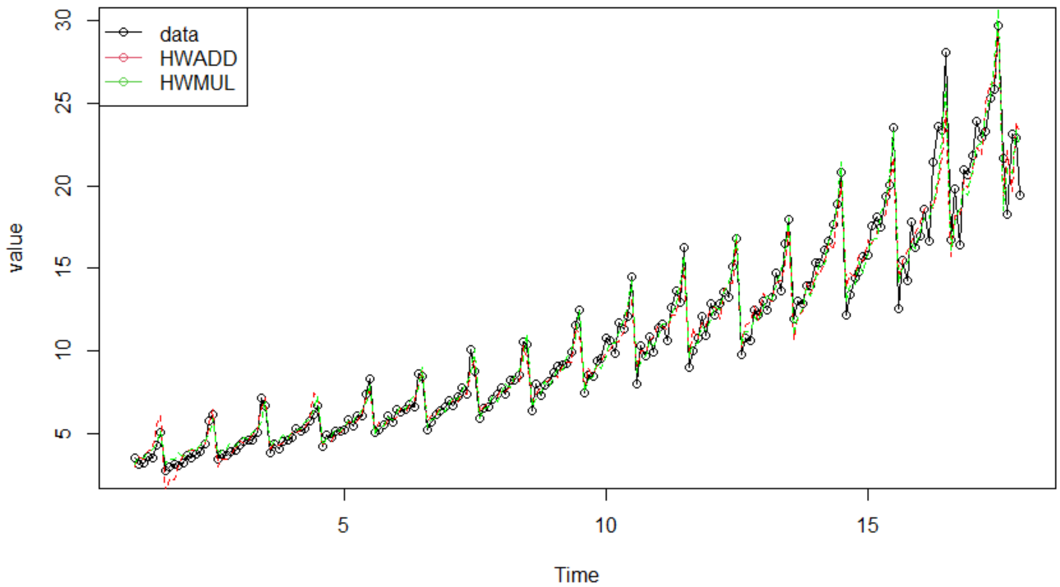

Figure 2. Two fitted models for the original time series.

It can be seen that both models fit the original data well. Now let's look at their prediction effects (See Figure 3).

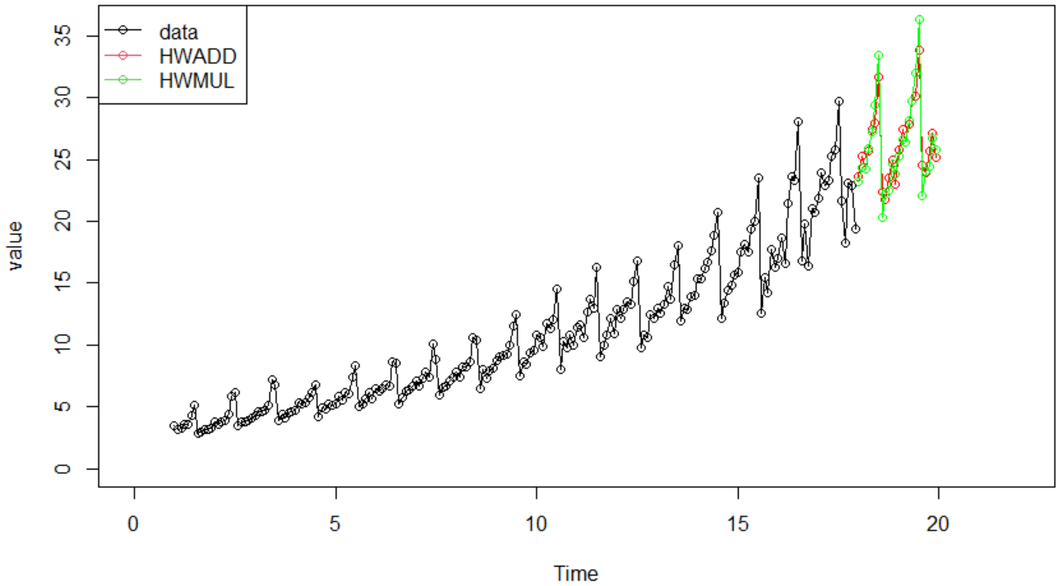

Figure 3. Two fitted models for the original time series.

It is difficult to judge which model is better only based on the above figure. After further checking the information of the model (See Figure 4), I choose the multiplicative model with better AIC, AICc and BIC.

```
Holt-Winters' additive method

Call:
 hw(y = x, seasonal = "additive")

  Smoothing parameters:
    alpha = 0.1466
    beta  = 0.0078
    gamma = 0.5041

  Initial states:
    l = 3.2104
    b = 0.0733
    s = -0.2957 -0.7423 -1.4985 -1.2063 -1.9913 2.7418
           1.9637 0.6541 0.4273 0.0018 0.2068 -0.2615

  sigma:  0.9761

     AIC     AICc      BIC
1092.350 1095.641 1148.758
```

```
Holt-Winters' multiplicative method

Call:
 hw(y = x, seasonal = "multiplicative")

  Smoothing parameters:
    alpha = 0.2303
    beta  = 0.0084
    gamma = 1e-04

  Initial states:
    l = 3.1538
    b = 0.146
    s = 0.91 0.9491 0.8737 0.8671 0.7982 1.325
           1.1754 1.0981 1.0503 0.9885 1.004 0.9606

  sigma:  0.0703

     AIC     AICc      BIC
925.1737 928.4640 981.5818
```

Figure 4. Model information for two fitted models.

2. The Fitted Model

The chosen model can be formulated as:

$$l_t = \alpha\left(\frac{x_t}{s_{t-m}}\right) + (1-\alpha)(l_{t-1} + g_{t-1})$$

$$g_t = \beta(l_t - l_{t-1}) + (1-\beta)g_{t-1}$$

$$s_t = \gamma\left(\frac{x_t}{l_{t-1} + g_{t-1}}\right) + (1+\gamma)s_{t-m}$$

$$\hat{x}_{t+h} = (l_t + g_t h)s_{t+h-m(k+1)}$$

where k is the integer part of $(h-1)/m$, $m = 12$ is the seasonal period here. All other needed parameters can be seen in Figure 4.

3. R code

```
x <- read.table("drug.txt", header = TRUE, sep = ",")
x <- ts(x['value'], frequency = 12)
library(fpp)
fit1 <- hw(x, seasonal="additive")
fit2 <- hw(x, seasonal="multiplicative")
plot(x, xlim=c(0, 22), ylim=c(0, 36))
plot(fit1)
plot(fit2)
lines(fitted(fit1), col="red", lty = 2)
lines(fitted(fit2), col="green", lty = 2)
lines(x, type="o")
lines(fit1$mean, type="o", col="red")
lines(fit2$mean, type="o", col="green")
legend ("topleft", lty=1, pch=1, col=1:3, c("data", "HWADD", "HWMUL"))
fit1$model
fit2$model
```