ISA 444: Business Forecasting

11 - Seasonal Decomposition and HW

Fadel M. Megahed

Associate Professor
Department of Information Systems and Analytics
Farmer School of Business
Miami University
Email: fmegahed@miamioh.edu
Office Hours: Click here to schedule an appointment

Spring 2021

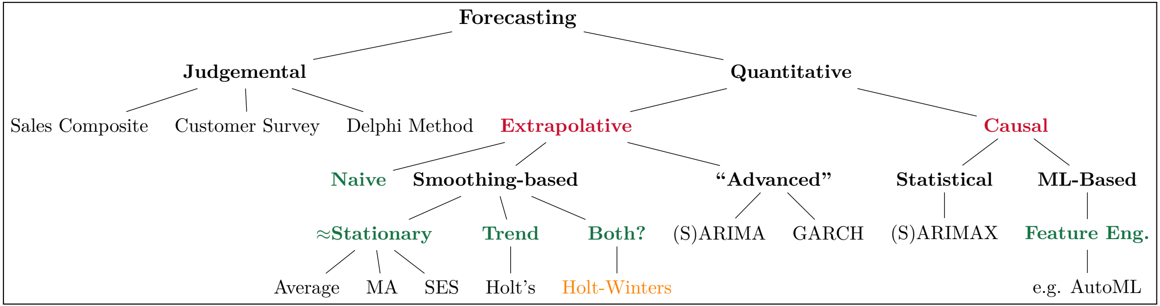
Outline

What we Covered Last Class

Main Learning Outcomes

- ☑ Recognize time series that are appropriate for linear exponential smoothing (LES).
- ☑ Use LES to forecast future observations of a time series.
- ☒ Explain when to use an additive vs. multiplicative model for a time series.
- ☒ Use classic decomposition methods to detrend and deseasonalize a time series.

Recap: A 10,000 Foot View of Forecasting Methods



A 10,000 foot view of forecasting techniques[1]

[1] An (incomplete) classification of forecasting techniques. Note that these focus on univariate time-series. Hence, they exclude popular approaches used in multivariate time series forecasting.

Recap: The LES Application to Multiple TS Example

```r
# Getting the crypto data (removed one of the coins to make the text fit here)
crypto = tq_get(c('BTC-USD', 'ETH-USD', 'ADA-USD', 'LINK-USD', 'ZIL-USD'), from = '2020-10-15')

crypto %<>% select(c(symbol, date, adjusted)) # keeping relevant columns
is_grouped_df(crypto) # checking if grouped by symbol; answer was FALSE (so we will group it)

crypto %<>% group_by(symbol) %>% mutate(adjustedLog = log(adjusted)) # to make data more linear

nestedCrypto = crypto %>% select(-c(date, adjusted)) %>% nest(data = adjustedLog)
nestedCrypto # printing to see the nested structure (uses list-column "innovation")

# map applies a function to every element of a list, when combined with mutate (we create new vars)
nestedCrypto %<>% mutate(data.ts = map(.x = data, .f = ts,
                                       start = c(2020, yday("2020-10-15")), freq = 366 ),
                         fitHolt = map(.x = data.ts, .f = holt, h = 30, alpha = 0.2, beta = 0.1),
                         accMetrics = map(.x = fitHolt, .f = accuracy),
                         MAPE = map(.x = accMetrics, .f = c(5)),
                         sweep = map(.x = fitHolt, .f = sw_sweep, fitted = T))
nestedCrypto # visually examining the output

unnestedCrypto = nestedCrypto %>% unnest(sweep) # expanding/unlisting the sweep var
```

Learning Objectives for Today's Class

Main Learning Outcomes

- Explain when to use an additive vs. multiplicative model for a time series.
- Use classic decomposition methods to detrend and deseasonalize a time series.
- Recognize time series that are appropriate for triple exponential smoothing (HW).
- Use HW to forecast future observations of a time series.

Outline

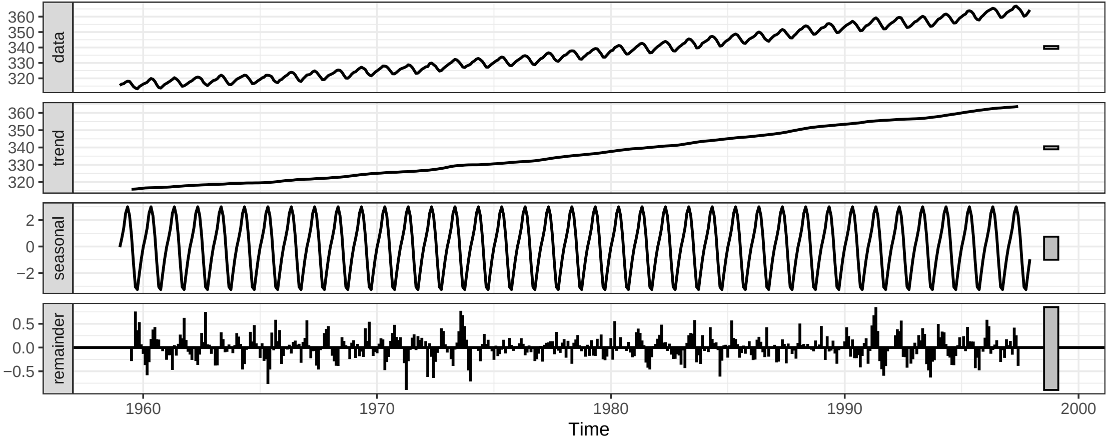
Definition and Basic Principles [1]

A time series may be made up of:

- **Trends (T)** - upward and downward movements
- **Seasonal (S) components** - regular, recurrent patterns that repeat at a fixed known duration (period)
- **Error (E) components** - irregular "noise" that is randomly distributed over time[2]

[2] A time series may also contain a cyclical component if it displays a somewhat periodic fluctuation, but the fluctuation has a periodicity of unknown duration, usually longer than a year.

Definition and Basic Principles [2]

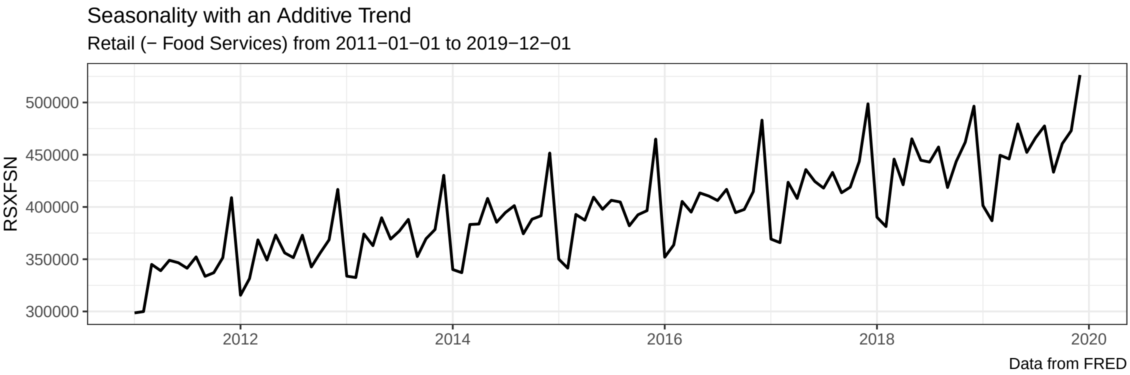


Based on C02 data in base R

Recall: Additive vs. Multiplicative Models [1]

An additive model is written as $Y = T + S + E$.

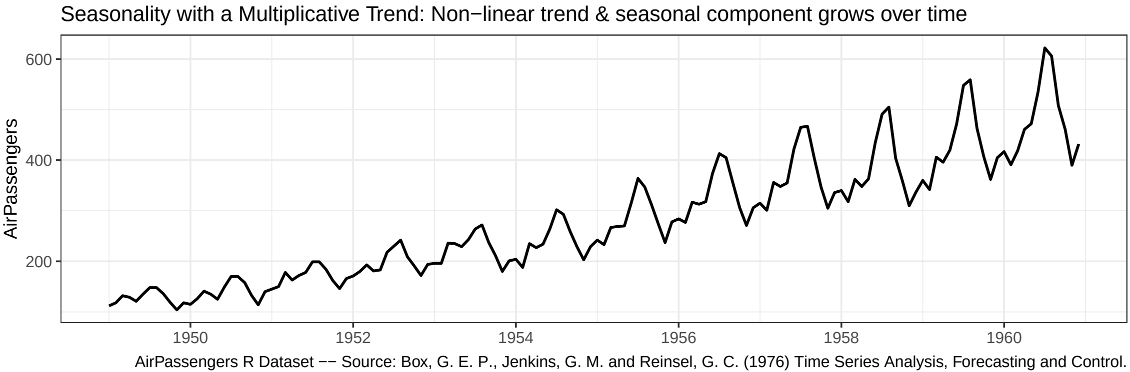
Definition: *An additive model is appropriate when the trend is approximately linear, and the seasonal components stays constant over time.*



Recall: Additive vs. Multiplicative Models [2]

A fully multiplicative model is written as Y = TSE.

Definition: *It is appropriate when the rate of change in the trend and/or the seasonal component and/or the variability in the error term increase or decrease over time.*



Some Comments

- When the trend and seasonal component are multiplied together, larger levels in the series will tend to exhibit larger peaks and troughs. When the error term is also multiplicative, the magnitude of the forecast errors will tend to rise and fall with the level of the series.[3]
- If the error variability is relatively constant over time, but the trend and/or seasonal components increase/decrease over time, a **mixed additive/multiplicative model**, $Y = TS + E$, may be more appropriate.
- An alternative to using a purely multiplicative model is to first transform the data using a logarithmic transformation.

$$
\begin{aligned}
Y &= TSE \\
\ln(Y) &= \ln(TSE) \\
&= \ln(T) + \ln(S) + \ln(E)
\end{aligned}
$$

[3] Slide is from Dr. Allison Jones-Farmer's lecture notes, Miami University, Spring 2020.

Outline

Background: Centered Moving Averages

Calculate the CMA(3), where you center the moving average in the middle of the moving window.

Q	Bike Sales	MA3
1.00	10.00	—-
2.00	31.00	—-
3.00	43.00	—-
4.00	16.00	—-
1.00	11.00	—-
2.00	33.00	—-
3.00	45.00	—-
4.00	17.00	—-
1.00	14.00	—-
2.00	36.00	—-
3.00	50.00	—-
4.00	21.00	—-
1.00	19.00	—-
2.00	41.00	—-
3.00	55.00	—-
4.00	25.00	—-

Decomposition Methods

Decomposition methods are used to "decompose" a time series into its components.
Decomposition methods are generally poor forecasting methods, but they work well for:

- exploring and visualizing time series data
- detrending and/or deseasonalizing data

Decomposition methods may be applied to multiplicative or additive time series.

Pure Decomposition Process for an Additive Time Series

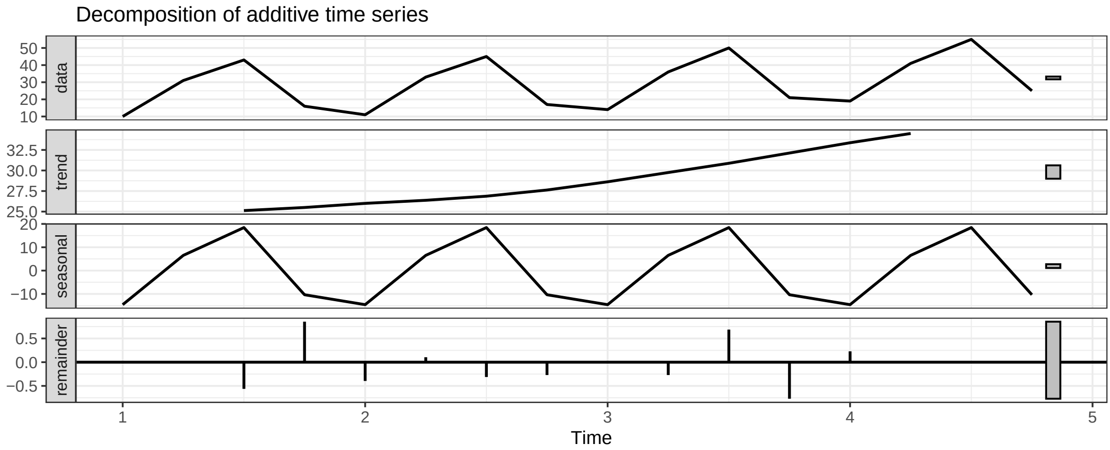
- **Estimate the trend** by calculating the centered moving average for a window of width K, denoted as CMA(K). Note you will lose $(K-1)/2$ observations at the beginning and end of the series if K is odd; suppose $K = 3$, so we lose one observation at the beginning and the end.
- **Detrend the series** by subtracting the CMA from the corresponding observations.
- **Estimate the initial seasonal factors** by calculating the average value of the detrended series for each quarter, month, day, etc. (depending on the season length).
- **Standardize the seasonal factors** by computing their averages and then setting the final seasonal factor for each season equal to the initial value minus the overall average.
- **Estimate the error term** by subtracting seasonal factor from the detrended series for each corresponding season.

Activity: Decomposing the BikeSalesR.xlsx

Based on the procedure described above, please use Excel/R to perform the aforementioned five steps. **You are expected to be able to do this once on your own.**

A Live Demo of Using R as an alternative

In class, we will use R to decompose the series and obtain the following plot



Notes on the `decompose()` in R

- The `decompose()` function in R uses a slightly different algorithm than your textbook presents.[4]
- The MA used to compute the trend estimate is a $2 \times m$ moving average. This means that for quarterly data, a 2×4 moving average is computed. First a MA(4) is computed, then a MA(2) of the MA(4) is computed. This is used to estimate the trend.
- The seasonal components are computed as usual and centered.

[4] Slide is from Dr. Allison Jones-Farmer's lecture notes, Miami University, Spring 2020.

Pure Decomposition Process for a Multiplicative Model

- **Estimate the trend** by calculating the centered moving average for a window of width K (i.e., CMA(K)). For now, let us assume that $k = 3$.
- **Detrend the series** dividing the observations $2, .., (n-1)$ from the their corresponding CMA(3).
- **Estimate the initial seasonal factors** by calculating the average value of the detrended series for each quarter, month, day, etc. (depending on the season length).
- **Standardize the seasonal factor** by computing their averages and then setting the final seasonal factor for each season equal to the initial value divided by the overall average.
- **Estimate the error term** by dividing the detrended series by the seasonal factor for each corresponding season.

Limitations to Decomposition

- Decomposition is widely used in practice but is not a good forecasting method.
- Decomposition methods are useful for visualizing your data and exploratory data analysis.
- Trend estimates are from moving averages and are not available for the first few and last few observations.
- Decomposition methods assume that the seasonal factors occur regularly from season to season over every period. This may not be true over the long run.
- Decomposition methods are not robust to unusual or spurious patterns that may occur in the data.

Because of these limitations, we need a better forecasting method for seasonal data![5]

[5] Slide is from Dr. Allison Jones-Farmer's lecture notes, Miami University, Spring 2020.

Outline

Definition and Basic Principles

If a time series has a linear trend with a local trend (β_1, growth rate) and a local seasonal pattern (SN_t) that may be changing over time, we can use the Holt-Winters exponential smoothing method for forecasting to accommodate the seasonal pattern.

The Holt-Winters method accommodates time series data with a **local level**, a **local trend**, and a **local seasonal pattern**, all of which are slowly changing over time. There are both additive and multiplicative versions of the Holt-Winters method.

Additive Holt-Winters Smoothing Method [1]

To compute the FORECAST, we will use three smoothing constants, α , to smooth the level, β , the smoothing constant to smooth the trend, and γ to smooth the seasonal pattern of length/frequency m (e.g. day-of-the-week pattern, $m = 7$; quarterly pattern, $m = 4$; monthly pattern, $m = 12$).

The estimate of the **level** is:

$$l_t = \alpha(y_t - sn_{t-L}) + (1 - \alpha)[l_{t-1} + b_{t-1}] \tag{1}$$

The estimate of the **trend** is:

$$b_t = \beta[l_t - l_{t-1}] + (1 - \beta)b_{t-1} \tag{2}$$

Additive Holt-Winters Smoothing Method [2]

The estimate of the **seasonal pattern** is:

$$sn_t = \gamma[y_t - l_t] + (1 - \gamma)sn_{t-L} \tag{3}$$

To estimate the **point forecast** for time $t + h$ time periods ahead made in time t:

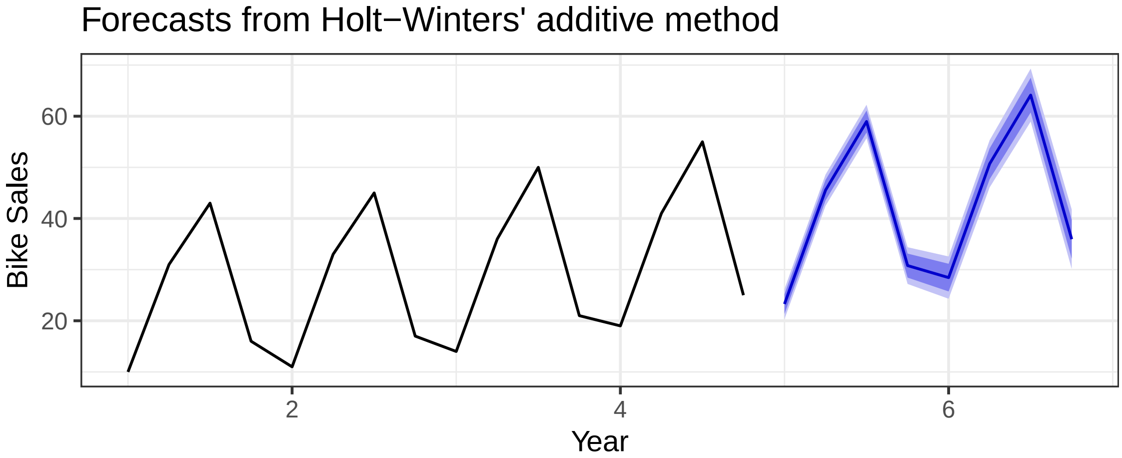
$$\hat{y}_{t+h}(t) = l_t + h \times b_t + sn_{t+h-L} \tag{4}$$

where sn_{t+h-L} is the most recent estimate of the seasonal pattern for the season corresponding to the time period $t + h$.

Comments on the Use of Software for Holt-Winters Method

- **Starting values:** We will need three sets of starting values; one for the Level, one for the Trend, and a set for m Seasonal Components. There are no two statistical packages that compute starting values in the same way! *Therefore, be comfortable on the fact that there will be some slight differences in the error values when compared to your textbook.*
- As we have done throughout the semester, we will be using R. The function used is titled `hw()`, which gets loaded from the forecast package (which we load when we run the command `pacman::p_load(fpp2)`).
- Details on the method used to compute starting values in the `hw()` function can be found in the Rstudio documentation by typing `?forecast::hw()` at the command prompt.

Live Demo: Holt Winters (Additive) on the BikeSales Data



Multiplicative Holt-Winters Smoothing Method [1]

To compute the FORECAST, we will use three smoothing constants, α , to smooth the level, β , the smoothing constant to smooth the trend, and γ to smooth the seasonal pattern of length/frequency m (e.g. day-of-the-week pattern, $m = 7$; quarterly pattern, $m = 4$; monthly pattern, $m = 12$).

The estimate of the **level** is:

$$l_t = \alpha(y_t/sn_{t-L}) + (1-\alpha)[l_{t-1} + b_{t-1}] \tag{5}$$

The estimate of the **trend** is:

$$b_t = \beta[l_t - l_{t-1}] + (1-\beta)b_{t-1} \tag{6}$$

Multiplicative Holt-Winters Smoothing Method [2]

The estimate of the **seasonal pattern** is:

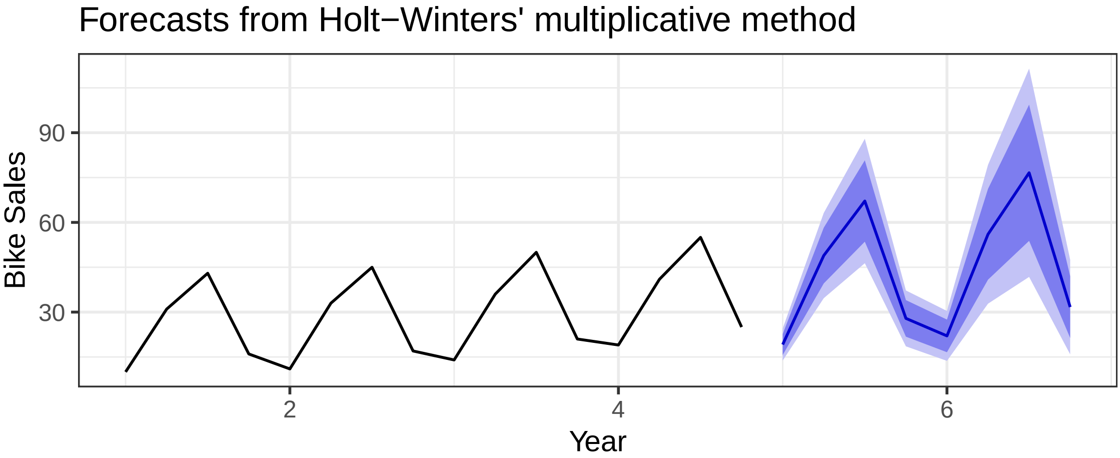
$$sn_t = \gamma[y_t/l_t] + (1-\gamma)sn_{t-L} \tag{7}$$

To estimate the **point forecast** for time $t+h$ time periods ahead made in time t:

$$\hat{y}_{t+h}(t) = (l_t + h \times b_t) \times sn_{t+h-L} \tag{8}$$

where sn_{t+h-L} is the most recent estimate of the seasonal pattern for the season corresponding to the time period $t+h$.

Live Demo: Holt Winters (Multiplicative) on BikeSales



Live Demo: Accuracy Comparison

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Additive HW	0.60	1.06	0.89	2.49	3.88	0.27	0.01
Multiplicative HW	-0.07	1.94	1.67	-1.75	7.86	0.50	0.18

Outline

Summary of Main Points

Main Learning Outcomes

- **Recognize time series that are appropriate for linear exponential smoothing (LES).**
- **Use LES to forecast future observations of a time series.**
- **Explain when to use an additive vs. multiplicative model for a time series.**
- **Use classic decomposition methods to detrend and deseasonalize a time series.**

Things to Do

- **Recommended:** Thoroughly read Chapter 4.1-4.4 and 4.6-4.7 of our textbook.
- Go through the slides, examples and make sure you have a good understanding of what we have covered.
- **Highly Recommended:** Go through the Week 06 Self-Paced Study Guide [**Will be published by 5PM on March 02, 2021**].
- **Required:** Complete the graded assignments 09-11. **Will be published on Canvas at noon on March 4, 2021.**

ISA 444: Business Forecasting

11 - Seasonal Decomposition and HW

Fadel M. Megahed

Associate Professor
Department of Information Systems and Analytics
Farmer School of Business
Miami University
Email: fmegahed@miamioh.edu
Office Hours: Click here to schedule an appointment

Spring 2021