

IENG/MANE 332 lecture notes

Reference: PRODUCTION, Planning, Control, and Integration by SIPPER & BULFIN

Chapter 4- Part 2: TIME SERIES METHODS (constant methods)

5. TIME SERIES METHODS

For short-term forecasting, time series methods are favored. A time series is simply a time-ordered list of historical data, the underlying assumption which is that history is a reasonable predictor of the future. There are several time series models and methods to choose from, including a constant, trend, or seasonal model, depending on the historical data and our understanding of the underlying process. For each model, there may be several forecasting methods available, including averages, moving averages, weighted moving average, exponential smoothing, regression, and even combinations of all them. Because we must recognize which model is appropriate for a given time series, we will discuss each model separately.

5.1. Constant Process

The Calgore Company is one of the largest producers of toothpaste in the United States. Almost 50 percent of their toothpaste is made at their New Jersey plant, with the rest of the production spread over five other plants dispersed across the country. As manager of toothpaste production, Ned Murphy is concerned about how much toothpaste he should produce next week. The actual sales figures for the last 50 weeks, obtained from the marketing department, are given in table 4.7. The first thing Ned does is plot the data; this plot is given in figure 4.7.

$$d_t = a + \varepsilon_t$$

In which represents the underlying constant of the process and ε_t the random noise, assumed to be normally distributed which mean zero and variance σ_ε^2 .

Many methods are used to forecast a constant process. We will discuss using the last data point, an average of all data points, an average of the most recent data, and averages that count all data points but give more weight to the more recent points.

We use some notations in this part:

t = an arbitrary time period

d_t = demand in period t

F_{t,t+k} = forecast made at time t for k periods ahead

5.1.1 SIMPLE METHODS. One of the simplest forecasting methods is to use the **last data point (LDP)** as the forecast for the next period.

Using LDP, the forecast for the next period will be the demand in this period. Notationally, it is

$$F_{t,t+1} = d_t$$

For Ned, the forecast for next week's demand would be 58, last week's demand. The forecast for k weeks in the future would also be

$$F_{t,t+k} = d_t$$

Because constant processes should have a constant mean and estimates of future demand should be independent of how far in the future we look.

The problem with LDP is the inherent random variation. If last week's demand was on the high side, the forecast will be too. If the next demand is high, the forecast will be good. However, for a constant model, we assume a normally distributed random component, and the next demand is just as likely to be low. In this case, LDP will give a bad forecast.

Table 4.7. Weekly toothpaste sales (in thousands of cases)

Week	Demand	Week	Demand	Week	Demand
1	56	18	55	35	52
2	46	19	52	36	48
3	53	20	52	37	50
4	50	21	44	38	49
5	50	22	47	39	52
6	52	23	57	40	48
7	46	24	45	41	47
8	53	25	48	42	48
9	55	26	55	43	44
10	46	27	50	44	43
11	53	28	42	45	50
12	45	29	50	46	57
13	50	30	57	47	46
14	49	31	51	48	44
15	48	32	54	49	52
16	43	33	54	50	58
17	47	34	51		

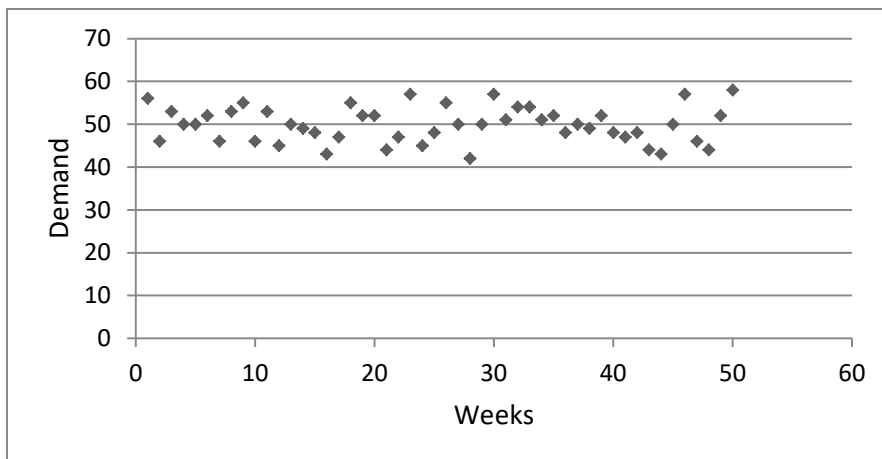


Figure 4.7. Weekly toothpaste sales plot

To overcome this problem, we could use an average of the past data, which will make the forecast less sensitive to random variations. Given t periods of data, the average at time t is

$$\bar{D}_t = \frac{1}{t} \sum_{i=1}^t d_i$$

The forecast for k periods in the future:

$$F_{t+1} = \bar{D}_t$$

For the data in table 4-7, we see that

$$\bar{D}_{50} = \frac{1}{50} \sum_{i=1}^t d_i = 49.88$$

Thus, the forecast for the next week (51) will be

$$F_{50,51} = \bar{D}_{50} = 49.88$$

Because units are in thousands of cases, the forecast is 49,880 cases.

To forecast for more than one period in the future, we would still use this number because we are using a constant model. Thus, the forecast for k periods in the future calculated at time T is

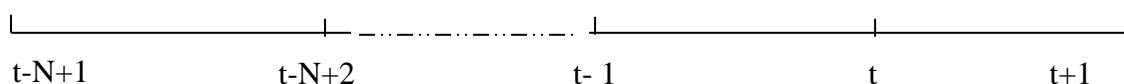
$$F_{t+k} = \bar{D}_t$$

The last data point and average forecasting methods could be considered extreme methods. LDP ignores all but the last data point, whereas the average treats very old data the same as the most recent. If the process is truly constant, an average is preferred because it captures the essence of the time series and tends to damp out random fluctuations. However, very few processes are constant over a long period of time. If the underlying processes changes, the last data point method will react to the change, but it will also react to random fluctuations. On the other hand, the average is slow to adjust to change but does not respond to random noise. Next we will examine some compromise methods.

5.1.2. MOVING AVERAGE (MA). Rather than take an average of all data points, we might choose to average only some of the more recent data. This method, called a moving average, is a compromise between the last data point and average methods. It averages recent data to reduce the effect of random fluctuations. Because only recent data is used to forecast, a moving average responds to a change in the underlying process more quickly. Let N be the number of periods we wish to consider in the moving average. We are currently at period t , the moving average is given by the sum of the last N data points, or mathematically.

$$F_{t,t+k} = \frac{1}{N} (d_{t-N+1} + d_{t-N+2} + \dots + d_t) = \frac{1}{N} \sum_{i=t-N+1}^t d_i$$

The timing of the data points in a moving average can be illustrated by the following time line:



At time t , the points $t-N+1, t-N+2, \dots, t$ are included in an N -period moving average. At time $t+1$ is added and the point $t-N+1$ is dropped from the calculation, so the average will include $t-N+2, t-N+3, \dots, t, t+1$.

In week 50, the five-week moving average of Calgore's sales would be

$$F_{50,51} = \frac{d_{46} + d_{47} + d_{48} + d_{49} + d_{50}}{5} = \frac{57 + 46 + 44 + 52 + 58}{5} = 51.4$$

Some more forecasts ($N=5$)

$$F_{40,51} = \frac{d_{36} + d_{37} + d_{38} + d_{39} + d_{40}}{5} = \frac{48 + 50 + 49 + 52 + 48}{5} = 49.4$$

$$F_{42,51} = \frac{d_{38} + d_{39} + d_{40} + d_{41} + d_{42}}{5} = \frac{49 + 52 + 48 + 47 + 48}{5} = 48.8$$

$$F_{46,51} = \frac{d_{42} + d_{43} + d_{44} + d_{45} + d_{46}}{5} = \frac{48 + 44 + 43 + 50 + 57}{5} = 48.4$$

Example: table 4-8 shows the demand for a specific product from 2000 to 2010, and 4-year moving average. Complete forecast values

Table 4.8. Yearly sales (in thousands of cases)

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Demand	255	252	265	248	246	260	258	261	250	258	255
Forecast					255	252.75	254.75	253	256.25	257.25	256.75

$$F_{2003,2004} = \frac{d_{2000} + d_{2001} + d_{2002} + d_{2003}}{4} = \frac{255 + 252 + 265 + 248}{4} = 255$$

$$F_{2004,2005} = \frac{d_{2001} + d_{2002} + d_{2003} + d_{2004}}{4} = \frac{252 + 265 + 248 + 246}{4} = 252.75$$

$$F_{2005,2006} = \frac{d_{2002} + d_{2003} + d_{2004} + d_{2005}}{4} = \frac{265 + 248 + 246 + 260}{4} = 254.75$$

$$F_{2006,2007} = \frac{d_{2003} + d_{2004} + d_{2005} + d_{2006}}{4} = \frac{248 + 246 + 260 + 258}{4} = 253$$

$$F_{2007,2008} = \frac{d_{2004} + d_{2005} + d_{2006} + d_{2007}}{4} = \frac{246 + 260 + 258 + 261}{4} = 256.25$$

$$F_{2008,2009} = \frac{d_{2005} + d_{2006} + d_{2007} + d_{2008}}{4} = \frac{260 + 258 + 261 + 250}{4} = 257.25$$

$$F_{2009,2010} = \frac{d_{2006} + d_{2007} + d_{2008} + d_{2009}}{4} = \frac{258 + 261 + 250 + 258}{4} = 256.75$$

Of course before calculating, we must construct scatter plot of data for checking behavior of variable during the time.

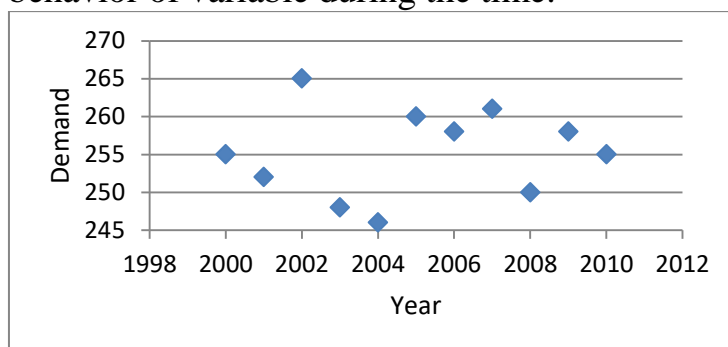


Figure 4.8. Yearly sales plot

We have constant process (there is no specific trend in data), so we can use moving average method for forecasting demand in future periods.

5.1.3. WEIGHTED MOVING AVERAGE.(WMA) The basic problem of simple moving average consists in assigning the same weights to all the recent data (demand) to calculate a forecast value, but it can be sometimes required that higher weights should be given on particular recent period's data.

This disadvantage can be overcome by using weighted moving average, and weighted moving average is also more suitable for the calculation of forecast values if there is a trend. The total weight is equal to 1, using the following equation.

$$F_{t,t+k} = (w_{t-N+1}d_{t-N+1} + w_{t-N+2}d_{t-N+2} + \dots + w_t d_t) = \sum_{i=t-N+1}^t w_i d_i$$

Example: From data of previous example, with 4-year weighted moving average. Complete forecast values

$$w_1 = 0.3, w_2 = 0.2, w_3 = 0.1, w_4 = 0.4$$

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Demand	255	252	265	248	246	260	258	261	250	258	255
Forecast					252.6	251.8	257.7	252.8	256	255.7	257.8

$$F_{2003,2004} = w_1 d_{2000} + w_2 d_{2001} + w_3 d_{2002} + w_4 d_{2003} \\ = 0.3 * 255 + 0.2 * 252 + 0.1 * 265 + 0.4 * 248 = 252.6$$

$$F_{2004,2005} = w_1 d_{2001} + w_2 d_{2002} + w_3 d_{2003} + w_4 d_{2004} \\ = 0.3 * 252 + 0.2 * 265 + 0.1 * 248 + 0.4 * 246 = 251.8$$

$$F_{2005,2006} = w_1 d_{2002} + w_2 d_{2003} + w_3 d_{2004} + w_4 d_{2005} \\ = 0.3 * 265 + 0.2 * 248 + 0.1 * 246 + 0.4 * 260 = 257.7$$

$$F_{2006,2007} = w_1 d_{2003} + w_2 d_{2004} + w_3 d_{2005} + w_4 d_{2006} \\ = 0.3 * 248 + 0.2 * 246 + 0.1 * 260 + 0.4 * 258 = 252.8$$

$$F_{2007,2008} = w_1 d_{2004} + w_2 d_{2005} + w_3 d_{2006} + w_4 d_{2007} \\ = 0.3 * 246 + 0.2 * 260 + 0.1 * 258 + 0.4 * 261 = 256$$

$$F_{2008,2009} = w_1 d_{2005} + w_2 d_{2006} + w_3 d_{2007} + w_4 d_{2008} \\ = 0.3 * 260 + 0.2 * 258 + 0.1 * 261 + 0.4 * 250 = 255.7$$

$$F_{2009,2010} = w_1 d_{2006} + w_2 d_{2007} + w_3 d_{2008} + w_4 d_{2009} \\ = 0.3 * 258 + 0.2 * 261 + 0.1 * 250 + 0.4 * 255 = 257.8$$

5.1.4. Determining the forecast accuracy for simple and weighted moving average methods.

In most cases the prediction accuracy is essential in choosing an appropriate forecasting method, whatever method of forecasting tends to be fairly inaccurate. In order to realize it, the actual values must be compared with estimated forecast.

The criteria used for evaluating the forecasting accuracy are given below.

- Mean Absolute Deviation (MAD) determines how the forecast accuracy has safer measure. To compute a MAD, the sum of absolute values of the forecast errors is divided by the number of forecasts.

$$MAD = \frac{\sum_{i=1}^t |d_i - F_i|}{t} = \frac{\sum_{i=1}^t |\varepsilon_i|}{t}$$

- Mean Absolute Percentage Error (MAPE) is calculated as

$$MAPE = 100 * \frac{\sum_{i=1}^t \frac{|e_i|}{d_i}}{t}$$

- Mean Square Error (MSE) is calculated as

$$MSE = \frac{\sum_{i=1}^t (d_i - F_i)^2}{t} = \frac{\sum_{i=1}^t (\varepsilon_i)^2}{t}$$

Table 4.9. Forecast comparing between two methods

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Demand	255	252	265	248	246	260	258	261	250	258	255
F-MA					255	252.75	254.75	253	256.25	257.25	256.75
F-WMA					252.6	251.8	257.7	252.8	256	255.7	257.8

Year	Demand	F-MA	$ \varepsilon_i $	$\frac{ e_i }{d_i}$	ε_i^2	F-WMA	$ \varepsilon_i $	$\frac{ e_i }{d_i}$	ε_i^2
2000	255								
2001	252								
2002	265								
2003	248								
2004	246	255	9	0.03659	81	252.6	6.6	0.02683	43.56
2005	260	252.75	7.25	0.02788	52.5625	251.8	8.2	0.03154	67.24
2006	258	254.75	3.25	0.01260	10.5625	257.7	0.3	0.00116	0.09
2007	261	253	8	0.03065	64	252.8	8.2	0.03142	67.24
2008	250	256.25	6.25	0.025	39.0625	256	6	0.024	36
2009	258	257.25	0.75	0.00291	0.5625	255.7	2.3	0.00891	5.29
2010	255	256.75	1.75	0.00686	3.0625	257.8	2.8	0.01098	7.84
		Σ	36.25	0.14249	250.8125	Σ	34.4	0.13484	227.26

$$MAD_{MA} = \frac{\sum_{i=1}^t |d_i - F_i|}{t} = \frac{\sum_{i=1}^t |\varepsilon_i|}{t} = \frac{36.25}{7} = 5.17$$

$$MAPE_{MA} = 100 * \frac{\sum_{i=1}^t \frac{|e_i|}{d_i}}{t} = \frac{0.142488}{7} = 2.0355$$

$$MSE_{MA} = \frac{\sum_{i=1}^t (d_i - F_i)^2}{t} = \frac{\sum_{i=1}^t (\varepsilon_i)^2}{t} = \frac{250.8125}{7} = 35.83$$

$$MAD_{WMA} = \frac{\sum_{i=1}^t |d_i - F_i|}{t} = \frac{\sum_{i=1}^t |\varepsilon_i|}{t} = \frac{34.4}{7} = 4.91$$

$$MAPE_{WMA} = 100 * \frac{\sum_{i=1}^t \frac{|e_i|}{d_i}}{t} = \frac{0.134843}{7} = 1.9263$$

$$MSE_{WMA} = \frac{\sum_{i=1}^t (d_i - F_i)^2}{t} = \frac{\sum_{i=1}^t (\varepsilon_i)^2}{t} = \frac{227.26}{7} = 32.466$$

Table shows the values of error for 4-week simple moving average, and 4-week weighted moving average.

	4-week simple moving average	4-week weighted moving average
<i>MAD</i>	5.17	4.91
<i>MAPE</i>	2.0355	1.9263
<i>MSE</i>	35.83	32.466

Forecasting using 4-week weighted moving average method is better than 4-week simple moving average since the weighted moving average provides smaller standard errors.

5.1.5. SIMPLE EXPONENTIAL SMOOTHING. Suppose we want to calculate an N period moving average but no longer know d_{T-N+1} , which is needed in the update formula. Our only choice is to estimate it.

$$F_{t,t+k} = \alpha d_t + (1 - \alpha)F_{t-1,t}$$

From the equation we see that α is the weight given to the most recent observation, so that a large weight will make the forecast more sensitive to the most recent data point. A smaller value will give more weight to an “average” value.

To implement exponential smoothing at time t , we need a value for $F_{t-1,t}$. Although there are many way to estimate $F_{t-1,t}$, we discuss about two of them:

- 1- Simple method: The simplest is to average several past data points (assume $N=5$).

Example:

Consider the data in table 4.7. Averaging the demand from weeks 45 to 49 gives

$$F_{49,50} = \frac{d_{45} + d_{46} + d_{47} + d_{48} + d_{49}}{5} = \frac{50 + 57 + 46 + 44 + 52}{5} = 49.8$$

$$F_{t,t+k} = \alpha d_t + (1 - \alpha)F_{t-1,t} \rightarrow F_{50,51} = 0.2d_{50} + (1 - 0.2)F_{49,50} \\ = 0.2 * 58 + 0.8 * 49.8 = 51.4$$

- 2- Regular method: Determination of the initial smoothed value to be set for the future forecast. The determination process becomes more difficult if we are dealing with many observations. Normally the initial smoothed value is taken equal to the first demand of observations.

Consider the data in table 4.7.

$$\begin{aligned} F_{50,51} &= 0.2d_{50} + (1 - 0.2)F_{49,50} = 0.2 * 58 + 0.8 * 48.71200511 = 50.56960409 \\ F_{49,50} &= 0.2d_{49} + (1 - 0.2)F_{48,49} = 0.2 * 52 + 0.8 * 47.89000639 = 48.71200511 \\ F_{48,49} &= 0.2d_{48} + (1 - 0.2)F_{47,48} = 0.2 * 44 + 0.8 * 48.86250799 = 47.89000639 \\ F_{47,48} &= 0.2d_{47} + (1 - 0.2)F_{46,47} = 0.2 * 46 + 0.8 * 49.57813499 = 48.86250799 \\ F_{46,47} &= 0.2d_{46} + (1 - 0.2)F_{45,46} = 0.2 * 57 + 0.8 * 47.72266874 = 49.57813499 \\ F_{45,46} &= 0.2d_{45} + (1 - 0.2)F_{44,45} = 0.2 * 50 + 0.8 * 47.15333592 = 47.72266874 \\ F_{44,45} &= 0.2d_{44} + (1 - 0.2)F_{43,44} = 0.2 * 43 + 0.8 * 48.1916699 = 47.15333592 \\ F_{43,44} &= 0.2d_{43} + (1 - 0.2)F_{42,43} = 0.2 * 44 + 0.8 * 49.23958738 = 48.1916699 \\ F_{42,43} &= 0.2d_{42} + (1 - 0.2)F_{41,42} = 0.2 * 48 + 0.8 * 49.54948422 = 49.23958738 \\ F_{41,42} &= 0.2d_{41} + (1 - 0.2)F_{40,41} = 0.2 * 47 + 0.8 * 50.18685527 = 49.54948422 \\ F_{40,41} &= 0.2d_{40} + (1 - 0.2)F_{39,40} = 0.2 * 48 + 0.8 * 50.73356909 = 50.18685527 \\ F_{39,40} &= 0.2d_{39} + (1 - 0.2)F_{38,39} = 0.2 * 52 + 0.8 * 50.41696136 = 50.73356909 \\ F_{38,39} &= 0.2d_{38} + (1 - 0.2)F_{37,38} = 0.2 * 49 + 0.8 * 50.7712017 = 50.41696136 \\ F_{37,38} &= 0.2d_{37} + (1 - 0.2)F_{36,37} = 0.2 * 50 + 0.8 * 50.96400212 = 50.7712017 \\ F_{36,37} &= 0.2d_{36} + (1 - 0.2)F_{35,36} = 0.2 * 48 + 0.8 * 51.70500265 = 50.96400212 \\ F_{35,36} &= 0.2d_{35} + (1 - 0.2)F_{34,35} = 0.2 * 52 + 0.8 * 51.63125331 = 51.70500265 \\ F_{34,35} &= 0.2d_{34} + (1 - 0.2)F_{33,34} = 0.2 * 51 + 0.8 * 51.78906664 = 51.63125331 \\ F_{33,34} &= 0.2d_{33} + (1 - 0.2)F_{32,33} = 0.2 * 54 + 0.8 * 51.2363333 = 51.78906664 \\ F_{32,33} &= 0.2d_{32} + (1 - 0.2)F_{31,32} = 0.2 * 54 + 0.8 * 50.54541662 = 51.2363333 \\ F_{31,32} &= 0.2d_{31} + (1 - 0.2)F_{30,31} = 0.2 * 51 + 0.8 * 50.43177077 = 50.54541662 \\ F_{30,31} &= 0.2d_{30} + (1 - 0.2)F_{29,30} = 0.2 * 57 + 0.8 * 48.78971346 = 50.43177077 \\ F_{29,30} &= 0.2d_{29} + (1 - 0.2)F_{28,29} = 0.2 * 50 + 0.8 * 48.48714182 = 48.78971346 \end{aligned}$$

$$\begin{aligned}
F_{28,29} &= 0.2d_{28} + (1 - 0.2)F_{27,28} = 0.2 * 42 + 0.8 * 50.10892728 = 48.48714182 \\
F_{27,28} &= 0.2d_{27} + (1 - 0.2)F_{26,27} = 0.2 * 50 + 0.8 * 50.1361591 = 50.10892728 \\
F_{26,27} &= 0.2d_{26} + (1 - 0.2)F_{25,26} = 0.2 * 55 + 0.8 * 48.92019888 = 50.1361591 \\
F_{25,26} &= 0.2d_{25} + (1 - 0.2)F_{24,25} = 0.2 * 48 + 0.8 * 49.1502486 = 48.92019888 \\
F_{24,25} &= 0.2d_{24} + (1 - 0.2)F_{23,24} = 0.2 * 45 + 0.8 * 50.18781075 = 49.1502486 \\
F_{23,24} &= 0.2d_{23} + (1 - 0.2)F_{22,23} = 0.2 * 57 + 0.8 * 48.48476344 = 50.18781075 \\
F_{22,23} &= 0.2d_{22} + (1 - 0.2)F_{21,22} = 0.2 * 47 + 0.8 * 48.8559543 = 48.48476344 \\
F_{21,22} &= 0.2d_{21} + (1 - 0.2)F_{20,21} = 0.2 * 44 + 0.8 * 50.06994287 = 48.8559543 \\
F_{20,21} &= 0.2d_{20} + (1 - 0.2)F_{19,20} = 0.2 * 52 + 0.8 * 49.58742859 = 50.06994287 \\
F_{19,20} &= 0.2d_{19} + (1 - 0.2)F_{18,19} = 0.2 * 52 + 0.8 * 48.98428574 = 49.58742859 \\
F_{18,19} &= 0.2d_{18} + (1 - 0.2)F_{17,18} = 0.2 * 55 + 0.8 * 47.48035717 = 48.98428574 \\
F_{17,18} &= 0.2d_{17} + (1 - 0.2)F_{16,17} = 0.2 * 47 + 0.8 * 47.60044646 = 47.48035717 \\
F_{16,17} &= 0.2d_{16} + (1 - 0.2)F_{15,16} = 0.2 * 43 + 0.8 * 48.75055808 = 47.60044646 \\
F_{15,16} &= 0.2d_{15} + (1 - 0.2)F_{14,15} = 0.2 * 48 + 0.8 * 48.9381976 = 48.75055808 \\
F_{14,15} &= 0.2d_{14} + (1 - 0.2)F_{13,14} = 0.2 * 49 + 0.8 * 48.922747 = 48.9381976 \\
F_{13,14} &= 0.2d_{13} + (1 - 0.2)F_{12,13} = 0.2 * 50 + 0.8 * 48.65343375 = 48.922747 \\
F_{12,13} &= 0.2d_{12} + (1 - 0.2)F_{11,12} = 0.2 * 45 + 0.8 * 49.56679219 = 48.65343375 \\
F_{11,12} &= 0.2d_{11} + (1 - 0.2)F_{10,11} = 0.2 * 53 + 0.8 * 48.70849024 = 49.56679219 \\
F_{10,11} &= 0.2d_{10} + (1 - 0.2)F_{9,10} = 0.2 * 46 + 0.8 * 49.3856128 = 48.70849024 \\
F_{9,10} &= 0.2d_9 + (1 - 0.2)F_{8,9} = 0.2 * 55 + 0.8 * 47.982016 = 49.3856128 \\
F_{8,9} &= 0.2d_8 + (1 - 0.2)F_{7,8} = 0.2 * 53 + 0.8 * 46.72752 = 47.982016 \\
F_{7,8} &= 0.2d_7 + (1 - 0.2)F_{6,7} = 0.2 * 46 + 0.8 * 46.5344 = 46.72752 \\
F_{6,7} &= 0.2d_6 + (1 - 0.2)F_{5,6} = 0.2 * 52 + 0.8 * 45.168 = 46.5344 \\
F_{5,6} &= 0.2d_5 + (1 - 0.2)F_{4,5} = 0.2 * 50 + 0.8 * 43.96 = 45.168 \\
F_{4,5} &= 0.2d_4 + (1 - 0.2)F_{3,4} = 0.2 * 50 + 0.8 * 53.7 = 43.96 \\
F_{3,4} &= 0.2d_3 + (1 - 0.2)F_{2,3} = 0.2 * 53 + 0.8 * 54 = 53.8 \\
F_{2,3} &= 0.2d_2 + (1 - 0.2)F_{1,2} = 0.2 * 46 + 0.8 * 56 = 54 \\
F_{1,2} &= d_1 = 56
\end{aligned}$$

Example: table represents the actual demand over the years 2012-2016. Forecast demand of 2017 with simple exponential smoothing method by two ways. Assume that $N=3$, $\alpha = 0.3$. Assume that real demand of 2017 will be 1425. Which method is better? Why?

year	2012	2013	2014	2015	2016
demand	1467	1500	1436	1395	1400

First method:

$$\begin{aligned}
F_{2015,2016} &= \frac{d_{2013} + d_{2014} + d_{2015}}{3} = \frac{1500 + 1436 + 1395}{3} = 1432 \\
F_{t,t+k} &= \alpha d_t + (1 - \alpha)F_{t-1,t} \rightarrow F_{2016,2017} = 0.3d_{2016} + (1 - 0.3)F_{2015,2016} \\
&= 0.3 * 1400 + 0.7 * 1432 = \mathbf{1422.4}
\end{aligned}$$

Second method:

$$\begin{aligned}
F_{t,t+k} &= \alpha d_t + (1 - \alpha)F_{t-1,t} \\
F_{2016,2017} &= 0.3d_{2016} + (1 - 0.3)F_{2015,2016} = 0.3 * 1400 + 0.7 * 1443.741 = \mathbf{1430.6187} \\
F_{2015,2016} &= 0.3d_{2015} + (1 - 0.3)F_{2014,2015} = 0.3 * 1395 + 0.7 * 1464.63 = 1443.741 \\
F_{2014,2015} &= 0.3d_{2014} + (1 - 0.3)F_{2013,2014} = 0.3 * 1436 + 0.7 * 1476.9 = 1464.63 \\
F_{2013,2014} &= 0.3d_{2013} + (1 - 0.3)F_{2012,2013} = 0.3 * 1500 + 0.7 * 1467 = 1476.9 \\
F_{2012,2013} &= d_{2012} = 1467
\end{aligned}$$

First method is better (because of lower error)