TIME SERIES

DECOMPOSITION METHODS FOR TIME SERIES ANALYSIS

The multiplicative decomposition model for a time series is

$$y_t = TR_t \times SN_t \times CL_t \times IR_t$$

where

$TR_t =$ the trend component

$SN_t =$ the seasonal component

$CL_t =$ the cyclical component

$IR_t =$ the irregular component

The multiplicative decomposition method can be used to obtain point estimates of the above 4 terms, in the following steps:

1. Compute MOVING AVERAGES (MA_t) and CENTERED MOVING AVERAGES (CMA_t) to eliminate seasonal and irregular components of the time series. For monthly data, seasons are months and length of moving average is 12; for quarterly data, this length is 4. Suppose we are given monthly data. Then if we are given data for 36 months, MA_t and CMA_t are calculated as follows:

$$MA_7 = \frac{\sum_{t=1}^{12} y_t}{12}, \quad MA_8 = \frac{\sum_{t=2}^{13} y_t}{12}, \ldots \quad MA_{31} = \frac{\sum_{t=25}^{36} y_t}{12}$$

$$CMA_7 = \frac{MA_7 + MA_8}{2}, \quad CMA_8 = \frac{MA_8 + MA_9}{2}, \ldots \quad MA_{30} = \frac{MA_{30} + MA_{31}}{2}$$

NOTE: $CMA_t = TR_t \times CL_t$

since the averaging in calculating the centered moving average should remove the seasonal and (short term) irregular components.

This implies

$CMA_t = TR_t \times CL_t$ and hence

$$SN_t \times IR_t = \frac{y_t}{TR_t \times CL_t} = \frac{y_t}{CMA_t}$$

2) We can calculate SN_t from $SN_t \times IR_t$ by grouping the values of $SN_t \times IR_t$ by months and calculating a monthly average $S\bar{N}_t$

3) Average SN_t are then normalized so they add to L = 12: the normalizing constant is

$$c = \frac{L}{\sum_{t=1}^{12} S\bar{N}_t}$$

Each $S\bar{N}_t$ is multiplied by the normalizing constant *c.*

4) We next calculate deseasonalized observations $d_t = \frac{y_t}{SN_t}$ and fit a trend equation to (t, d_t) data:

$TR_t = d_t = \beta_0 + \beta_1 t$ if d_t appears to be linearly related to t.

5) We now have estimates of SN_t and TR_t , and

$$CL_t \times IR_t = \frac{y_t}{TR_t \times SN_t}$$

6) Calculate CL_t for monthly or quarterly data by calculating the 3-period moving average

$$CL_t = \frac{CL_{t-1}IR_{t-1} + CL_t IR_t + CL_{t+1}IR_{t+1}}{3}$$

7) Next calculate the irregular component IR_t by the formula

$$IR_t = \frac{CL_t \times IR_t}{CL_t}$$

8) We now have estimates of the 4 components. In case the irregular component IR_t *is close to 1 for all t,* the point forecast for y_t *is*

$$y_t = \begin{cases} TR_t \times SN_t \times CL_t & \text{if a well-defined cycle exists and can be predicted} \\ TR_t \times SN_t & \text{if a well-defined cycle DOES NOT exist} \end{cases}$$

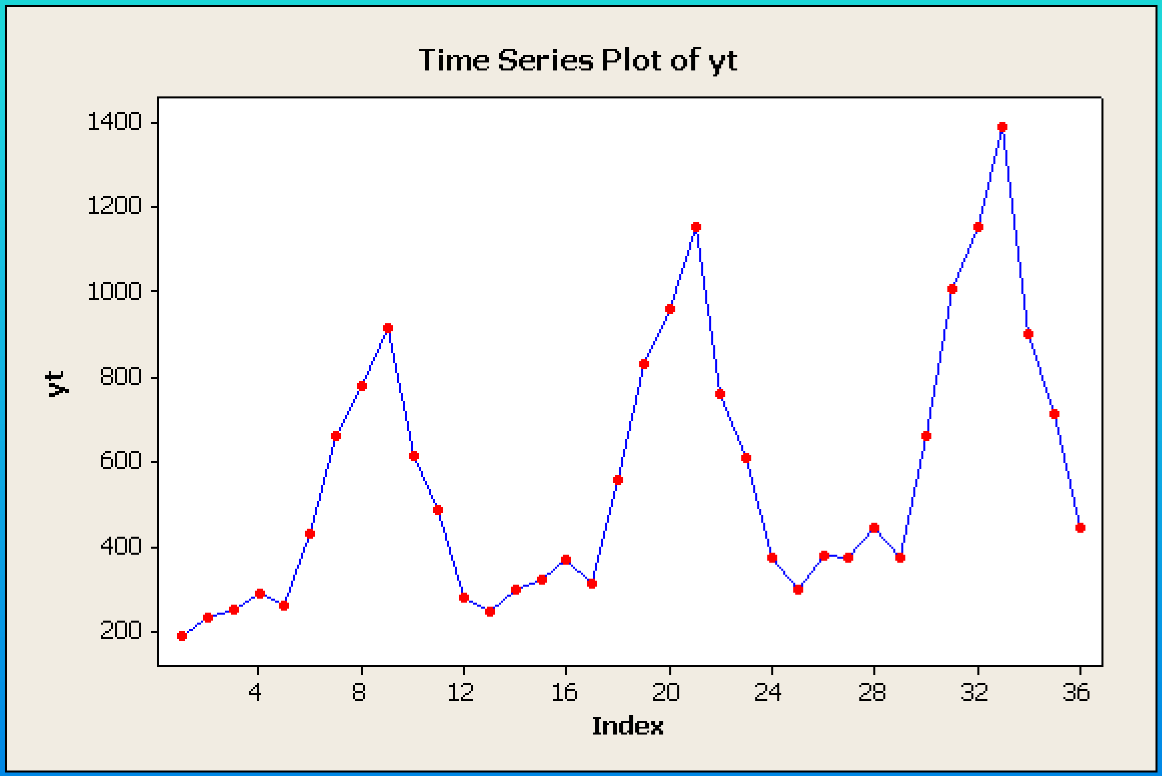
9) There is no theoretically correct prediction interval for y_t but an approximate formula is

$\hat{y}_t \pm B_t[100(1-\alpha)]$ where $B_t[100(1-\alpha)]$ is the error bound in a 100(1-α)% prediction interval $TR_t \pm B_t[100(1-\alpha)]$ for the deseasonalized observation $d_t = \beta_0 + \beta_1 t$.

Example 1 (Example 7.1/page 328, , Bowerman, O'Connell, Koehler, 4-th Edition – Forecasting, Time Series, and Regression)

y_t = sale of Tasty Cola for month t

month	t	yt	month	t	yt	month	t	yt
1	1	189	1	13	244	1	25	298
2	2	229	2	14	296	2	26	378
3	3	249	3	15	319	3	27	373
4	4	289	4	16	370	4	28	443
5	5	260	5	17	313	5	29	374
6	6	431	6	18	556	6	30	660
7	7	660	7	19	831	7	31	1004
8	8	777	8	20	960	8	32	1153
9	9	915	9	21	1152	9	33	1388
10	10	613	10	22	759	10	34	904
11	11	485	11	23	607	11	35	715
12	12	277	12	24	371	12	36	441



Time Series Plot of yt
yt
1400
1200
1000
800
600
400
200
4
8
12
16
20
24
28
32
36
Index

Year 1

month	t	y	MA12	CMA12 =Tr x Cl	Sn x Ir = yt/(Tr x Cl)	Sn	d=y/Sn
1	1	189				0.492878	383.4619
2	2	229				0.595136	384.7857
3	3	249				0.594957	418.5173
4	4	289				0.679406	425.3718
5	5	260				0.563812	461.1466
6	6	431				0.984748	437.6755
7	7	660	447.8333	450.125	1.466259372	1.465826	450.258
8	8	777	452.4167	455.2083	1.706910755	1.691517	459.3509
9	9	915	458	460.9167	1.985174471	1.988409	460.1669
10	10	613	463.8333	467.2083	1.312048515	1.306178	469.3083
11	11	485	470.5833	472.7917	1.025821803	1.027936	471.8193
12	12	277	475	480.2083	0.576832972	0.599571	461.9969

Year 2

month	t	y	MA12	CMA12 =Tr x Cl	Sn x Ir = yt/(Tr x Cl)	Sn	d=y/Sn
1	13	244	485.4167	492.5417	0.49539	0.492878	495.0513
2	14	296	499.6667	507.2917	0.583491	0.595136	497.3649
3	15	319	514.9167	524.7917	0.60786	0.594957	536.1728
4	16	370	534.6667	540.75	0.684235	0.679406	544.5936
5	17	313	546.8333	551.9167	0.567115	0.563812	555.1495
6	18	556	557	560.9167	0.991235	0.984748	564.6116
7	19	831	564.8333	567.0833	1.465393	1.465826	566.9158
8	20	960	569.3333	572.75	1.676124	1.691517	567.5378
9	21	1152	576.1667	578.4167	1.991644	1.988409	579.3576
10	22	759	580.6667	583.7083	1.300307	1.306178	581.0848
11	23	607	586.75	589.2917	1.03005	1.027936	590.5037
12	24	371	591.8333	596.1667	0.622309	0.599571	618.7757

Year 3

month	t	y	MA12	CMA12 =Tr x Cl	Sn x Ir = yt/(Tr x Cl)	Sn	d=y/Sn
1	25	298	600.5	607.7083	0.490367	0.492878	604.6119
2	26	378	614.9167	622.9583	0.606782	0.595136	635.1485
3	27	373	631	640.8333	0.582055	0.594957	626.9356
4	28	443	650.6667	656.7083	0.674576	0.679406	652.0405
5	29	374	662.75	667.25	0.56051	0.563812	663.3416
6	30	660	671.75	674.6667	0.978261	0.984748	670.2224
7	31	1004	677.5833			1.465826	684.9379
8	32	1153				1.691517	681.6365
9	33	1388				1.988409	698.0455
10	34	904				1.306178	692.0957
11	35	715				1.027936	695.5686
12	36	441				0.599571	735.5258

Year 1

TR	yhat	Cl x Ir	Cl	Ir
=380.4+9.498 t	=Tr x Sn	=y/(TR x Sn)	=(MA3 for Cl x Ir)	=Cl x Ir/Cl
389.898	192.1722	0.983492829		
399.396	237.6951	0.96341902	0.990148959	0.97300412
408.894	243.2745	1.023535028	1.001212147	1.02229586
418.392	284.2579	1.016682392	1.039313213	0.97822522
427.89	241.2495	1.077722218	1.031687343	1.04462096
437.388	430.7168	1.000657417	1.028641734	0.97279488
446.886	655.0572	1.007545567	1.004901281	1.00263139
456.384	771.9815	1.006500858	1.000593023	1.00590433
465.882	926.364	0.987732644	0.993820371	0.99387442
475.38	620.9308	0.987227611	0.982676078	1.00463177
484.878	498.4236	0.97306798	0.964933588	1.00843
494.376	296.4136	0.934505173	0.963354489	0.97005327

Year 2

TR =380.4+9.498 t	yhat =Tr x Sn	Cl x Ir =y/(TR x Sn)	Cl =(MA3 for Cl x Ir)	Ir =Cl x Ir/Cl
503.874	248.3485041	0.982490315	0.961938408	1.021365
513.372	305.5263934	0.968819737	0.992250651	0.976386
522.87	311.0853962	1.025441901	1.005742085	1.019587
532.368	361.6938391	1.022964618	1.024306969	0.99869
541.866	305.5105951	1.024514387	1.023835327	1.000663
551.364	542.9544462	1.024026977	1.019778348	1.004166
560.862	822.1262326	1.01079368	1.009957504	1.000828
570.36	964.773841	0.995051855	1.001660866	0.993402
579.858	1152.994961	0.999137064	0.993384861	1.005791
589.356	769.8036824	0.985965665	0.990386303	0.995536
598.854	615.5835867	0.98605618	0.996385374	0.989633
608.352	364.7502685	1.017134275	0.993921441	1.023355

Year 3

TR =380.4+9.498 t	yhat =Tr x Sn	Cl x Ir =y/(TR x Sn)	Cl =(MA3 for Cl x Ir)	Ir =Cl x Ir/Cl
617.85	304.5248	0.978573867	1.002714052	0.97592516
627.348	373.3577	1.012434012	0.991815397	1.02078876
636.846	378.8963	0.984438312	1.001895237	0.9825761
646.344	439.1298	1.008813385	1.001562251	1.00723982
655.842	369.7716	1.011435056	1.009195556	1.00221909
665.34	655.1921	1.007338225	1.011246582	0.99613511
674.838	989.1952	1.014966466	1.006120005	1.00879265
684.336	1157.566	0.996055324	1.005697212	0.99041273
693.834	1379.626	1.006069846	0.995383123	1.01073629
703.332	918.6766	0.984024198	0.988626242	0.99534501
712.83	732.7436	0.975784681	0.992693359	0.98296687
722.328	433.087	1.018271197		

Forecasting

Since the multiplicative decomposition for tasty cola data shows no well-defined cycle or irregular component. Hence

$y_t = TR_t \times SN_t$

We next compute the forecasts and approximate 95% prediction intervals for y_t for the 4-th year.

An approximate 95% prediction interval for y_t *is*

$$\hat{y}_t \pm B_t[100(1-\alpha)] \text{ where}$$

$$B_t[100(1-\alpha)] = \text{ error bound for the deseasonalized trend line } d_t = \beta_0 + \beta_1 t$$

$$= 2 \times t_{n-2,\frac{\alpha}{2}} s \times \sqrt{1 + d(t_0)}$$

where

$$d(t_0) = \frac{1}{n} + \frac{(t_0 - \overline{t})^2}{S_{tt}}$$

$$S_{tt} = \sum_{i=1}^{n} (t_i - \overline{t})^2$$

t0	d(t0)	TRt	SNt	Yhat_t	B(t0)	PL95	PU95
37	0.049352	731.256	0.49288	360.42	25.9544	334.47	386.37
38	0.051748	740.745	0.59514	440.84	25.984	414.86	466.83
39	0.054269	750.234	0.59496	446.36	26.0152	420.34	472.37
40	0.056917	759.723	0.67941	516.16	26.0478	490.11	542.21
41	0.05969	769.212	0.56381	433.69	26.082	407.61	459.77
42	0.06259	778.701	0.98475	766.82	26.1176	740.71	792.94
43	0.065616	788.19	1.46583	1155.35	26.1548	1129.19	1181.5
44	0.068768	797.679	1.69152	1349.29	26.1934	1323.09	1375.48
45	0.072045	807.168	1.98841	1604.98	26.2336	1578.75	1631.21
46	0.075449	816.657	1.30618	1066.7	26.2752	1040.42	1092.97
47	0.078979	826.146	1.02794	849.23	26.3183	822.91	875.54
48	0.082636	835.635	0.59957	501.02	26.3628	474.66	527.39

ADDITIVE DECOMPOSITION

$$y_t = TR_t + SN_t + CL_t + IR_t$$

where

$TR_t =$ the trend component

$SN_t =$ the seasonal component

$CL_t =$ the cyclical component

$IR_t =$ the irregular component

Steps of additive decomposition are similar to those of multiplicative decomposition.

1) Calculate CMA values.

$$CMA_t = TR_t + CL_t$$

since the averaging in calculating the centered moving average should remove the seasonal and (short term) irregular components.

This implies

$CMA_t = TR_t + CL_t$ and hence

$$SN_t + IR_t = y_t - (TR_t + CL_t) = y_t - CMA_t$$

2) Calculate $S\bar{N}_t$ from $SN_t + Ir_t$ *values, and normalize so the* normalized sum equals 0.

$$SN_t = S\bar{N}_t - \frac{\sum_{t=1}^{L} S\bar{N}_t}{L}$$

3) Calculate deseasonalized observations

$$d_t = y_t - SN_t$$

4) Fit a trend model to $d_t = y_t - SN_t$

5) $CL_t + IR_t = y_t - TR_t - SN_t$

6) Calculate CL_t for monthly or quarterly data by calculating the 3-period moving average

$$CL_t = \frac{CL_{t-1}IR_{t-1} + CL_t IR_t + CL_{t+1}IR_{t+1}}{3}$$

and then $IR_t = (CL_t + IR_t) - CL_t$

$$\hat{y}_t \pm B_t[100(1-\alpha)] \text{ where}$$

$$B_t[100(1-\alpha)] = \text{ error bound for the deseasonalized trend line } d_t = \beta_0 + \beta_1 t$$

$$= 2 \times t_{n-2,\frac{\alpha}{2}} s \times \sqrt{1 + d(t_0)}$$

where

$$d(t_0) = \frac{1}{n} + \frac{(t_0 - \overline{t})^2}{S_{tt}}$$

$$S_{tt} = \sum_{i=1}^{n} (t_i - \overline{t})^2$$