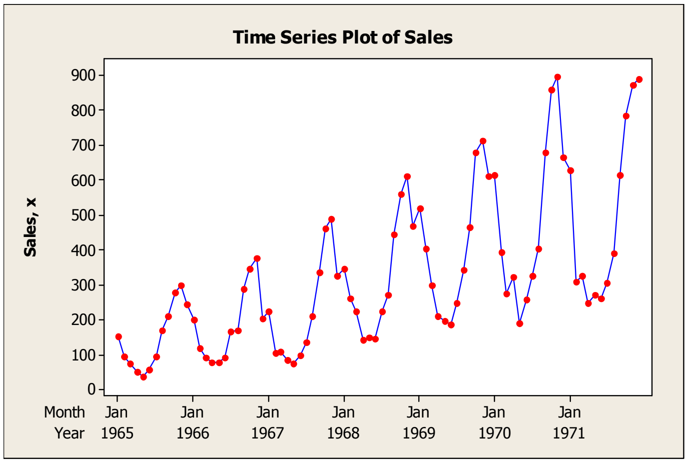
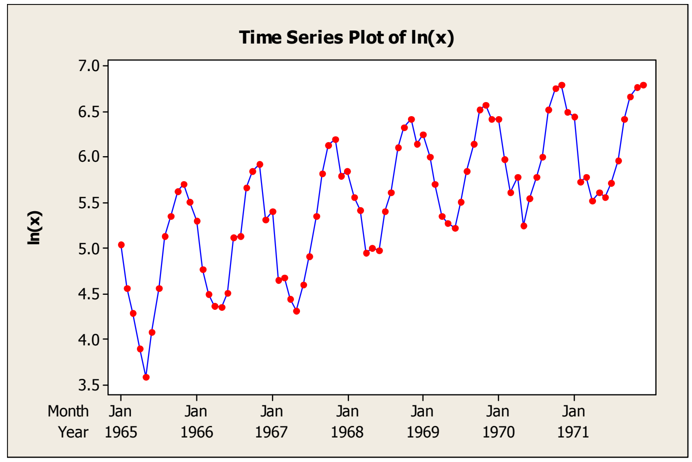
Minitab Project Report – Computer Lab 4

4.1 Sales of an industrial heater, January 1965 till December 1971



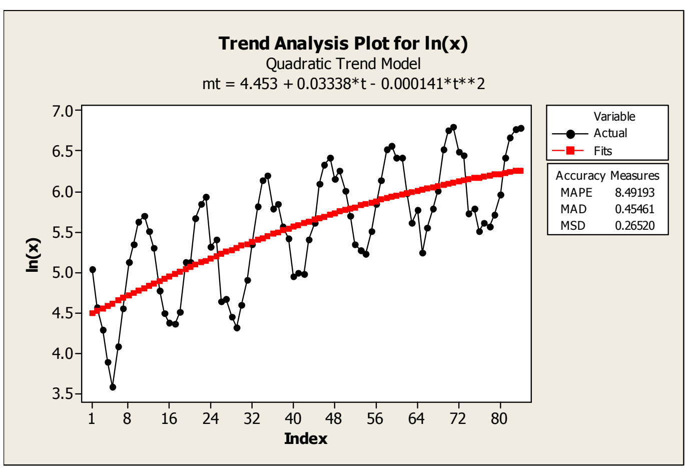


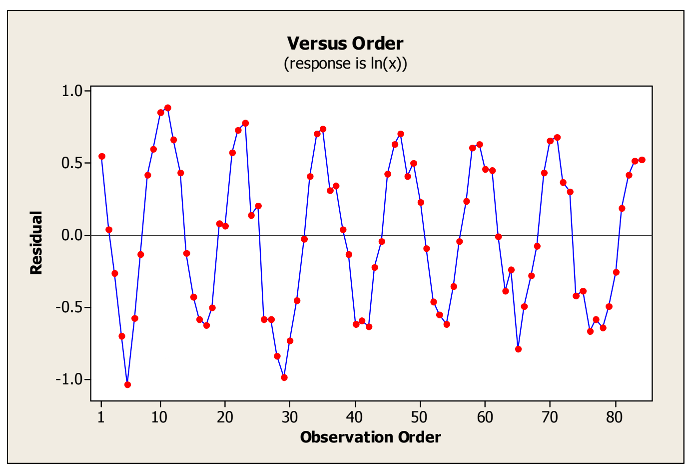
The data show seasonality, an increasing trend and also increasing variance. The log transformed data stabilize the variance.

The model we will be fitting is

$$\ln(x_t) = m_t + s_t + y_t,$$

where m_t denotes trend, s_t denotes a seasonality effect and y_t denotes a random noise.





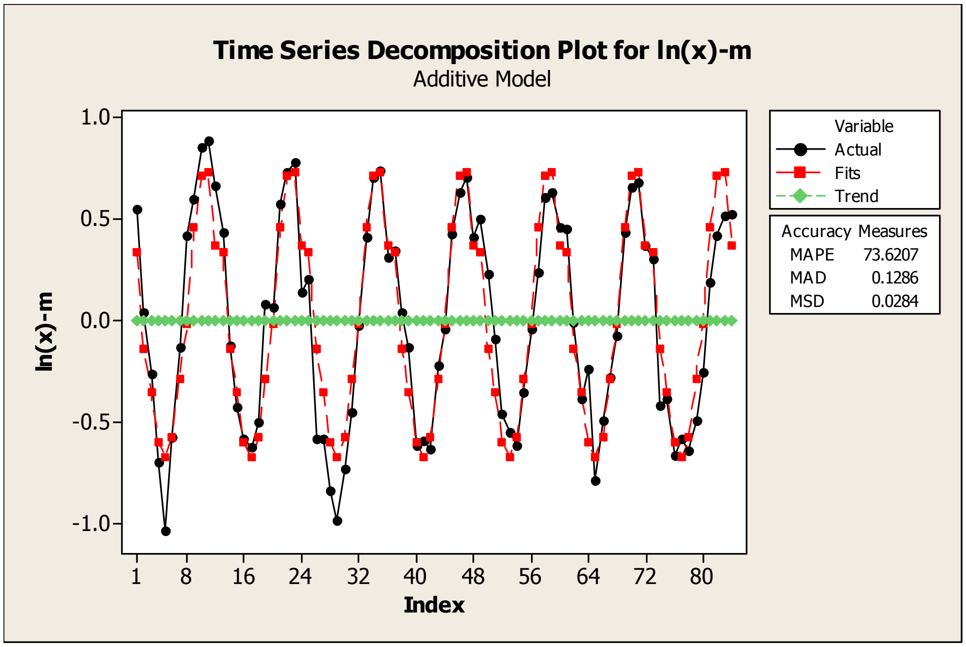
The quadratic model,

$$m_t = 4.453 + 0.03338\, t - 0.000141\, t^2,$$

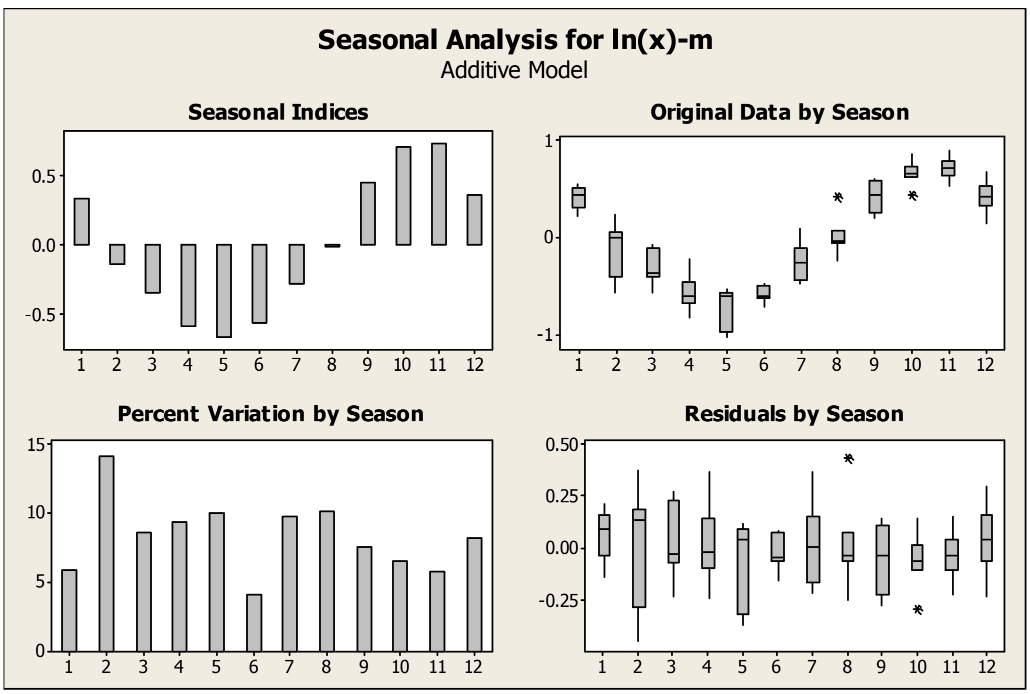
fits the trend well. The residuals, which are in fact the detrended data, oscillate about zero, but still show the seasonality effects and noise.

Time Series Decomposition for ln(x)-m

```
Additive Model

Data      ln(x)-m
Length    84
NMissing  0

Seasonal Indices

Period       Index
     1    0.341767
     2   -0.140364
     3   -0.352853
     4   -0.599822
     5   -0.668700
     6   -0.574700
     7   -0.281610
     8   -0.009815
     9    0.461350
    10    0.714300
    11    0.738105
    12    0.372342
```

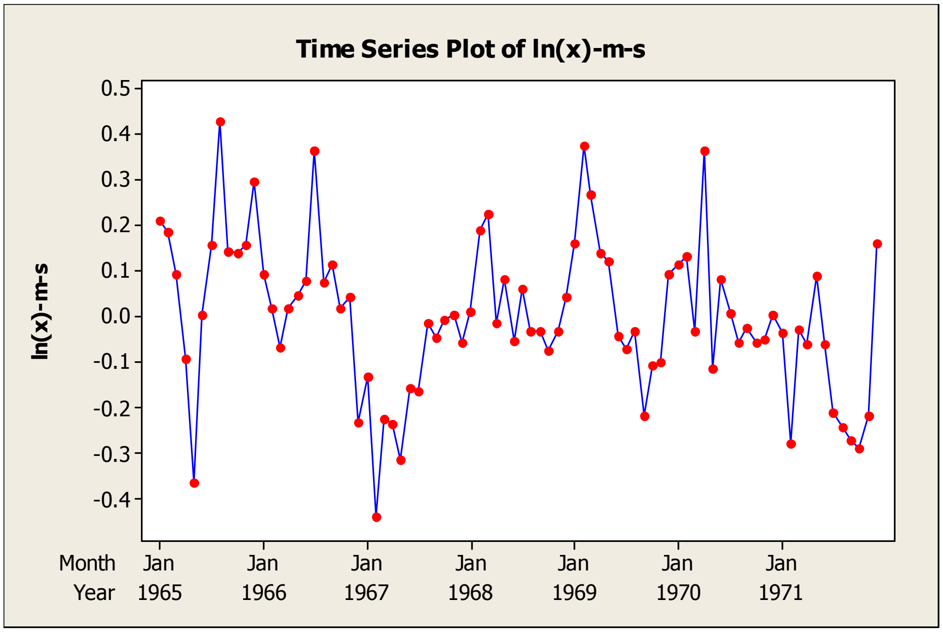


The seasonal indices show that the lowest sales are, on average, in May, while the highest sales are in November. The graph obviously shows no trend (detrended data). Hence, the Fits here are in fact the seasonal effects.



The Seasonal Analysis of the detrended data shows that although May has the lowest seasonality effect, there is a large variability in the effect as well as in May's residual.
The medians of the seasonality effects of April, May and June are very similar; hence all these three months may be comparably bad for the sales.

There is also a large variation in the effect of February as well as in February's residuals.



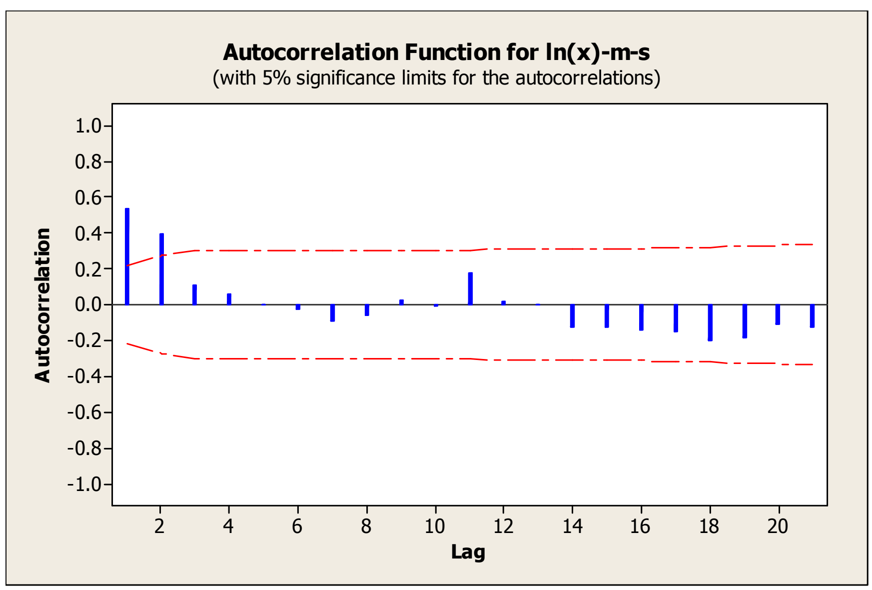
The residuals of the decomposition are the detrended and deseasonalized data, which represent a realization of the random noise

$$y_t = \ln(x_t) - m_t - s_t.$$

They oscillate abut zero, there is neither a clear trend nor seasonality, but there are some local trends.

Autocorrelation Function: ln(x)-m-s

Lag	ACF	T	LBQ
1	0.535627	4.91	24.97
2	0.396072	2.89	38.79
3	0.106325	0.71	39.80
4	0.059756	0.40	40.12
5	-0.002977	-0.02	40.12
6	-0.022045	-0.15	40.17
7	-0.094506	-0.63	41.00
8	-0.061470	-0.40	41.36
9	0.022981	0.15	41.41
10	-0.005449	-0.04	41.42
11	0.178639	1.17	44.58
12	0.016259	0.11	44.60



The sample autocorrelation function suggests that there are two non-zero values and then the autocorrelations become non-significant. It looks like the sample ACF cuts off after lag two. This suggests an MA(2) model for the random noise y_t.

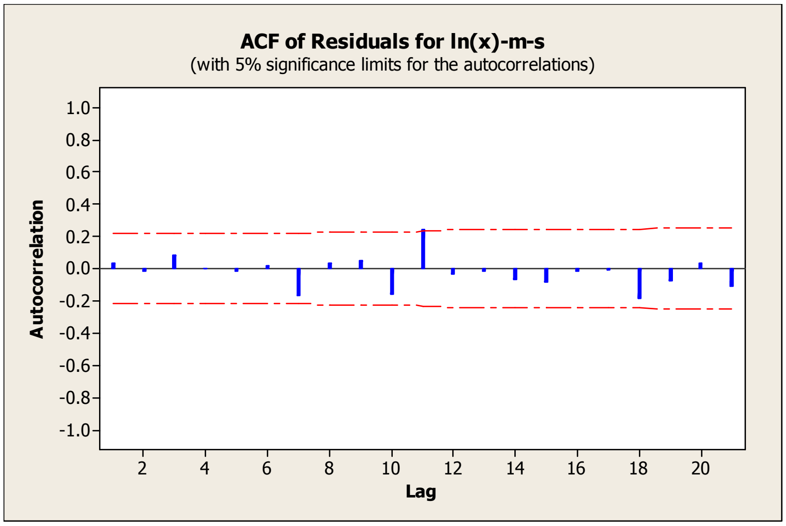
ARIMA Model: ln(x)-m-s

Final Estimates of Parameters

Type		Coef	SE Coef	T	P
MA	1	-0.4765	0.0961	-4.96	0.000
MA	2	-0.5494	0.0960	-5.72	0.000

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	12.2	21.1	26.6	38.9
DF	10	22	34	46
P-Value	0.275	0.517	0.813	0.761



The numerical output, as well as the sample ACF for the noise of the fitted MA(2) model suggest a good model fit. Both model parameters (θ_1 and θ_2) are strongly significant, the p-values of the Ljung-Box test are large, that is none of the groups of the residuals are correlated and so the residuals may represent a White Noise variable.

The fitted MA(2) is

$$y_t = z_t + 0.4765\, z_{t-1} + 0.5494\, z_{t-2},$$

where z_t is a representation of a White Noise random variable.

Hence, the final model for the transformed sales of the industrial heater can be written as

$$\ln(x_t) = 4.453 + 0.03338\, t - 0.000141\, t^2 + s_t + z_t + 0.4765\, z_{t-1} + 0.5494\, z_{t-2},$$

where s_t denotes the fitted seasonal effect, such that $s_t = s_{t+12}$.