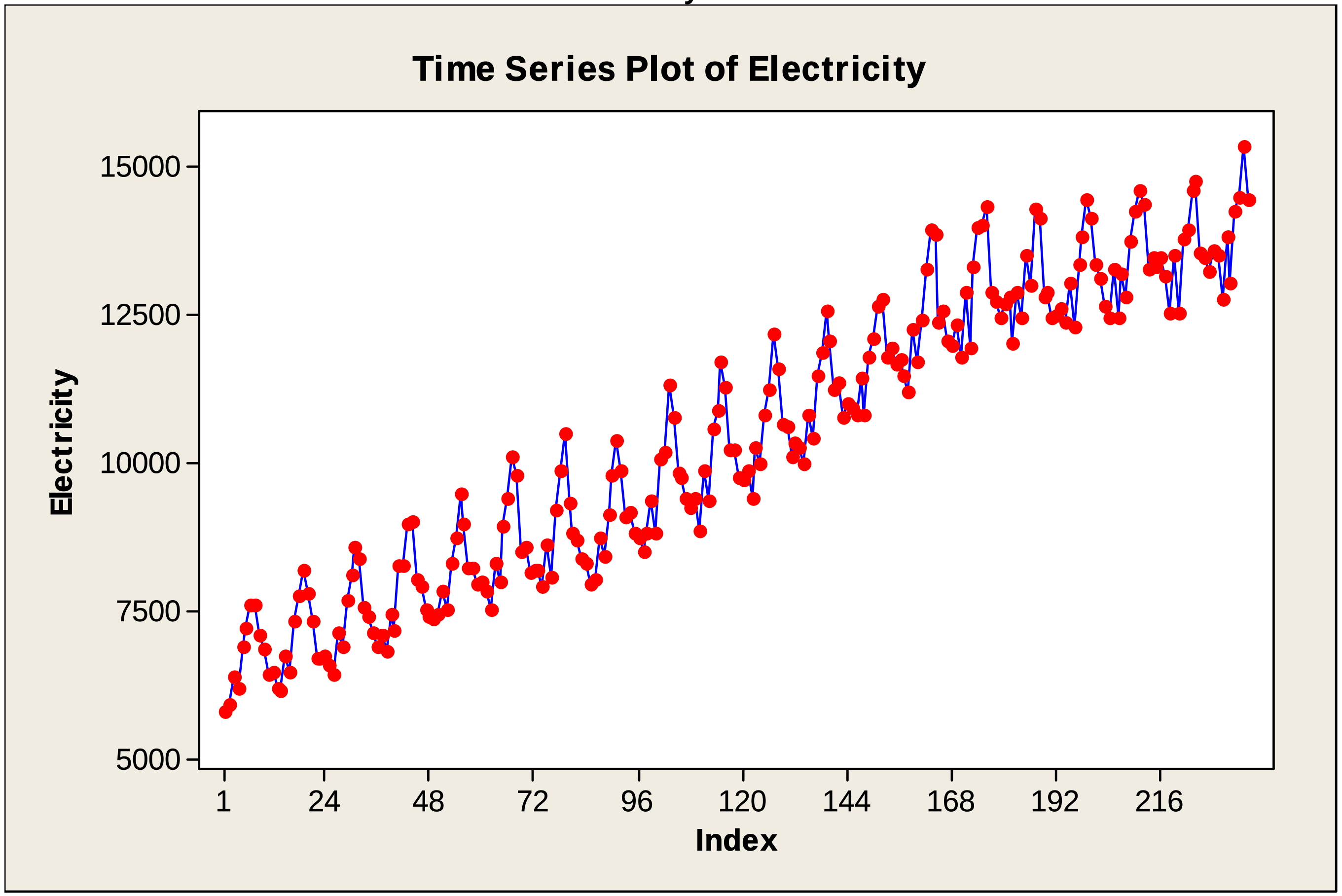
Minitab Project Report Assignment 2

2.1.1 Plot of the data

Australian Electricity Production Data



These are monthly observations; the period length is 12. The data indicate clear trend and seasonality.

Over the years the production steadily increases, although it slowed down slightly in last four years.

The seasonal effects are additive as they seem similar over all periods (variability does not increase with time).

May-August show higher production then the rest of the year; a clear peak is in July.

2.1.2 Small Trend Method

We assume an additive model

$X_t = m_t + s_t + Y_t$ (1)

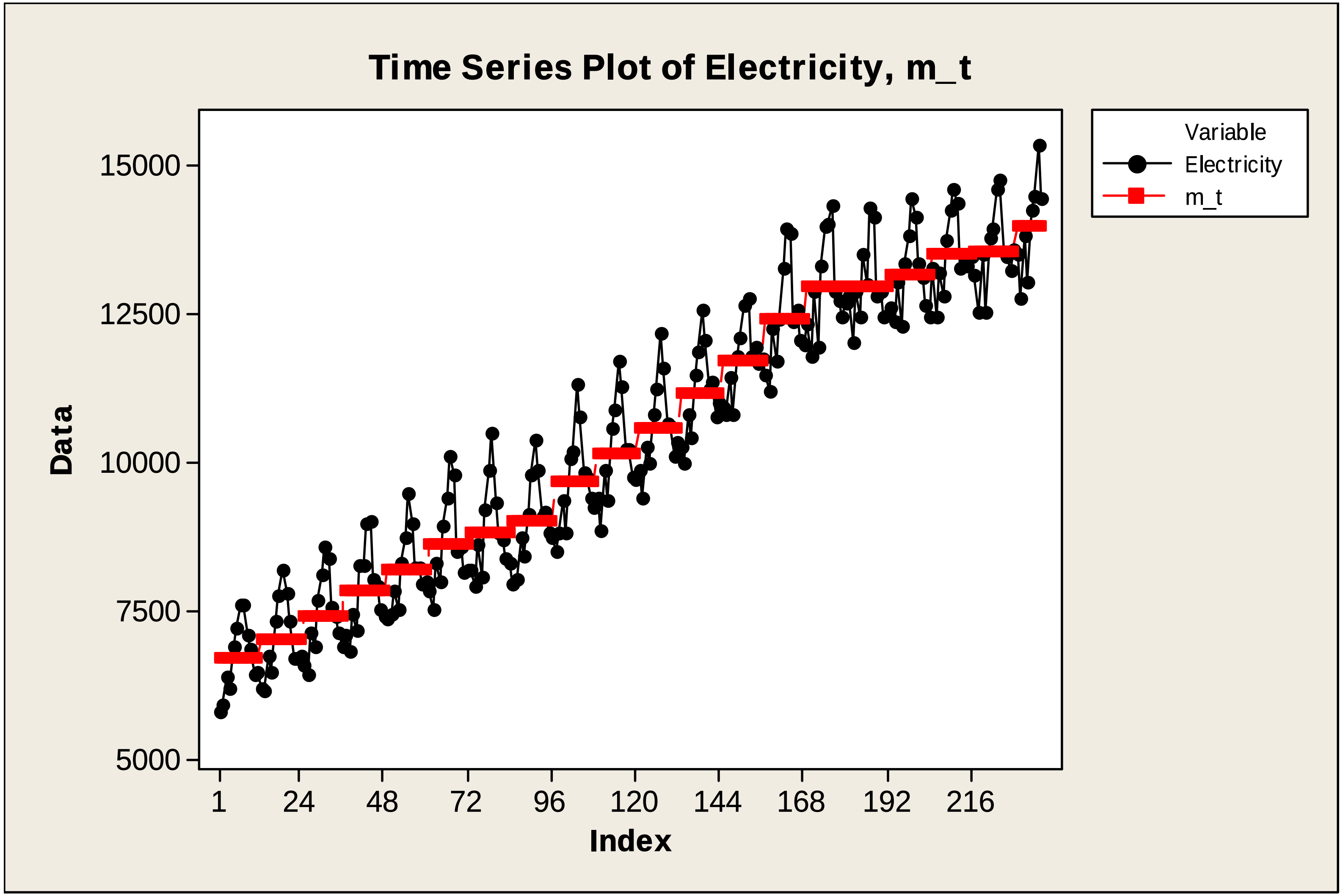
where

$E(Y_t) = 0$ (2)

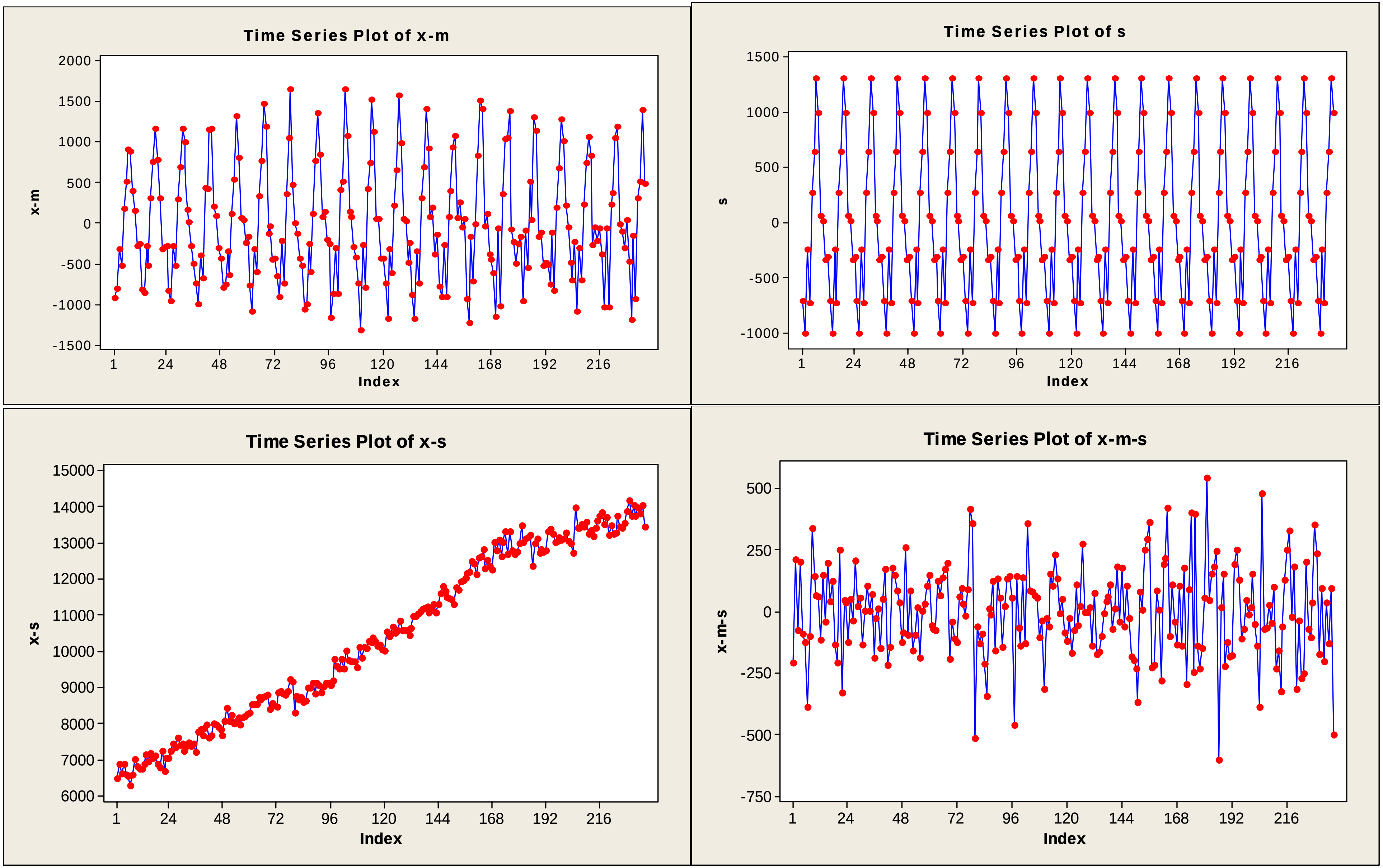
$s_t = s_{t-d}$ (the seasonality effect is assumed to be the same for the same seasons) (3)

$\Sigma s_k = 0$ (the seasonality effects over one period sum to zero) (4)

The small trend method relies on the assumption that the trend is constant over each period and it is estimated by an average of the observations for all seasons in the period.
Each seasonal component is estimated by an average of detrended data over observations for the given season in all periods.



Time Series Plot of Electricity, m_t
Variable
Electricity
m_t
15000
12500
10000
7500
5000
Data
1
24
48
72
96
120
144
168
192
216
Index



Time Series Plot of x-m
x-m
Index
Time Series Plot of s
s
Index
Time Series Plot of x-s
x-s
Index
Time Series Plot of x-m-s
x-m-s
Index

Comments:

The detrended data set (x-m, where m is the mean over 12 month of a year) is still showing clear seasonality effects and noise effects, but no trend. The values fluctuate about zero.

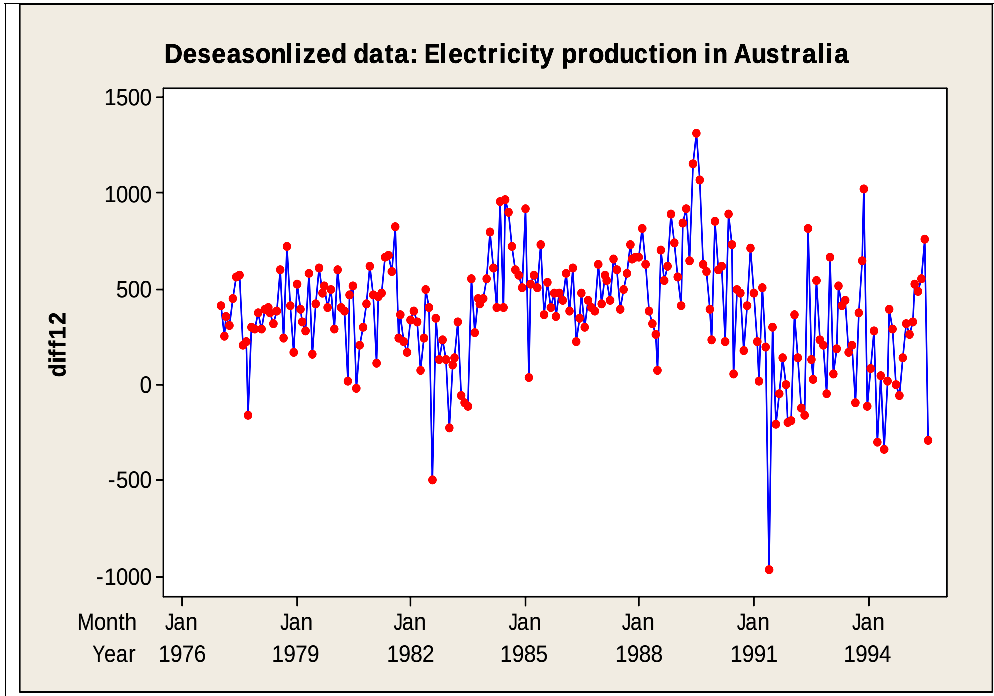
The seasonality effects (s) show a peak in electricity production in July and the lowest value in February.

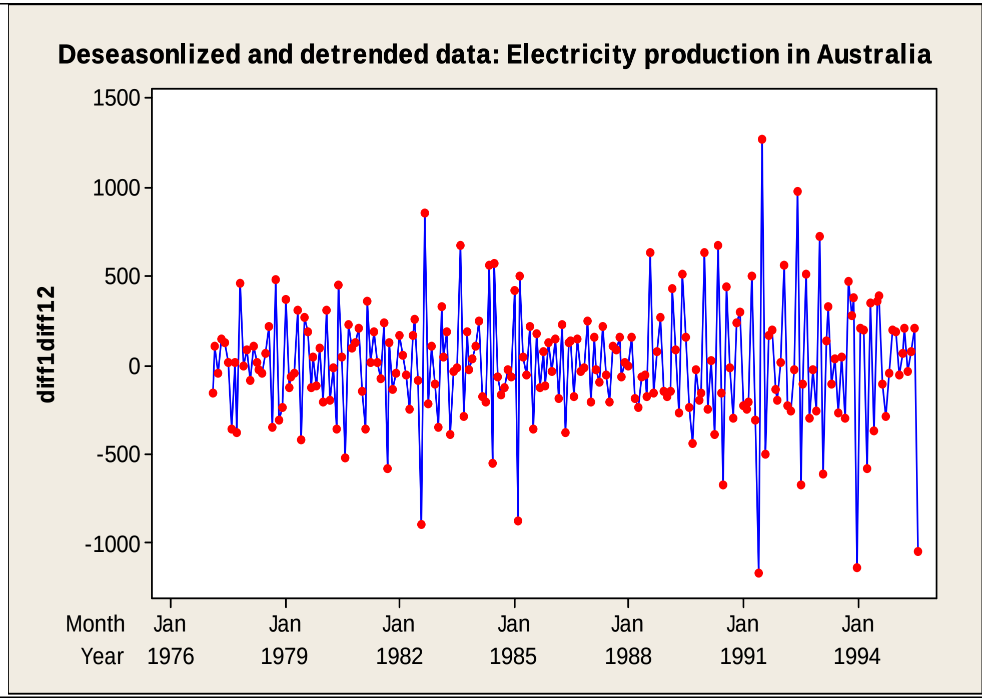
The de-seasonalized data show clear upward trend in the electricity production over the years.

The residuals are free of the trend and seasonality effects. The TS looks stationary with mean zero.

Difference Method

Assuming model (1) we may remove the trend or seasonality effects by differencing.
Differencing by the lag equal to the length of period (lag = d) may remove seasonality effect, assuming (3).
Differencing by a small lag may remove trend (for example taking lag = 1 if the trend is approximately linear).

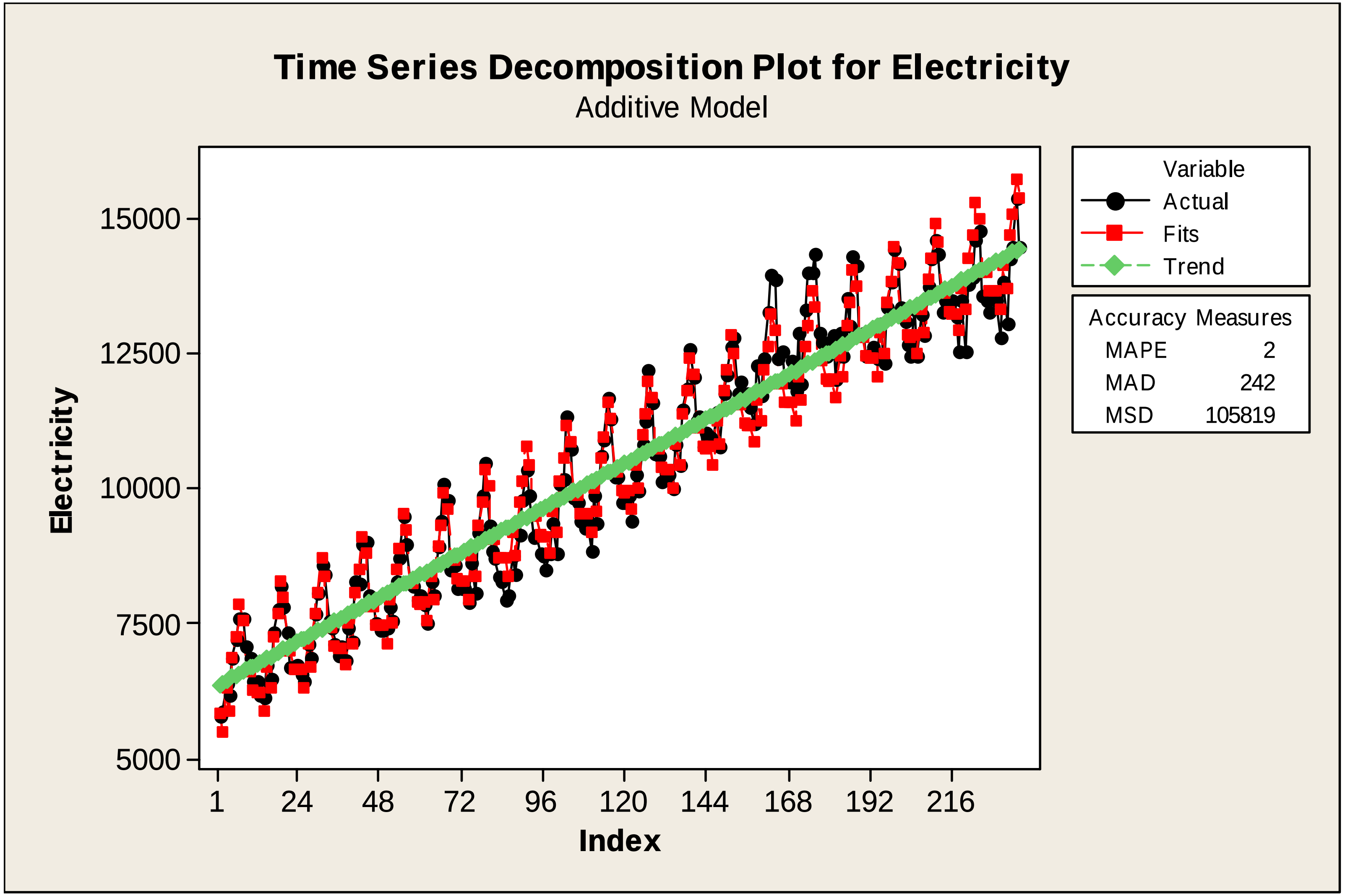


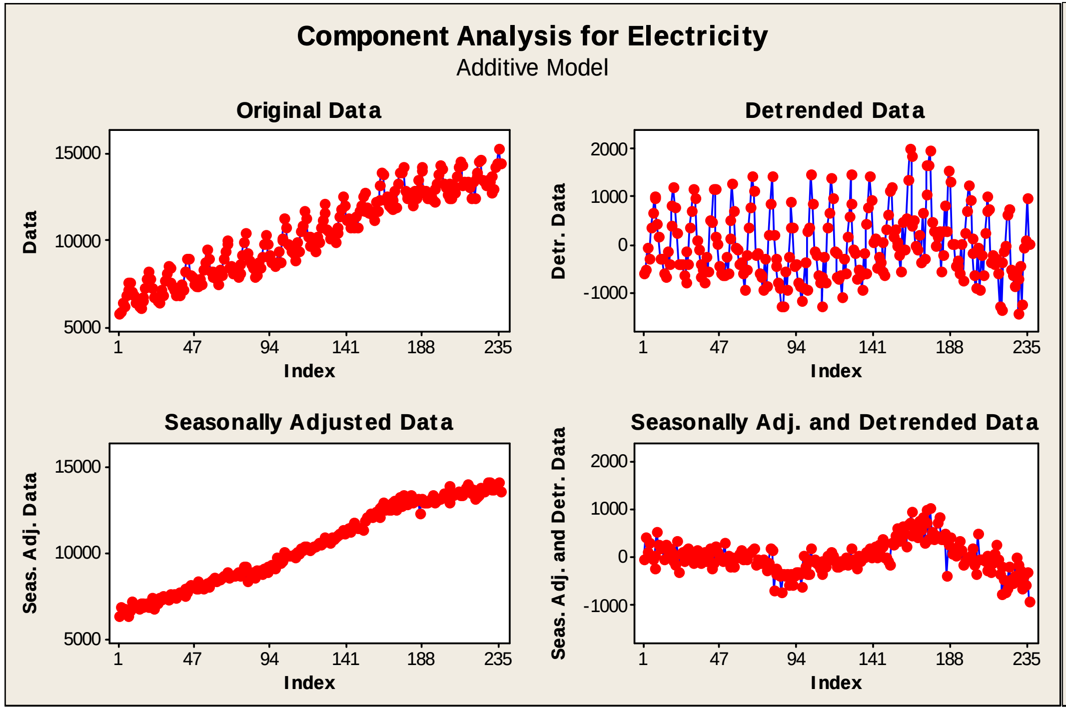


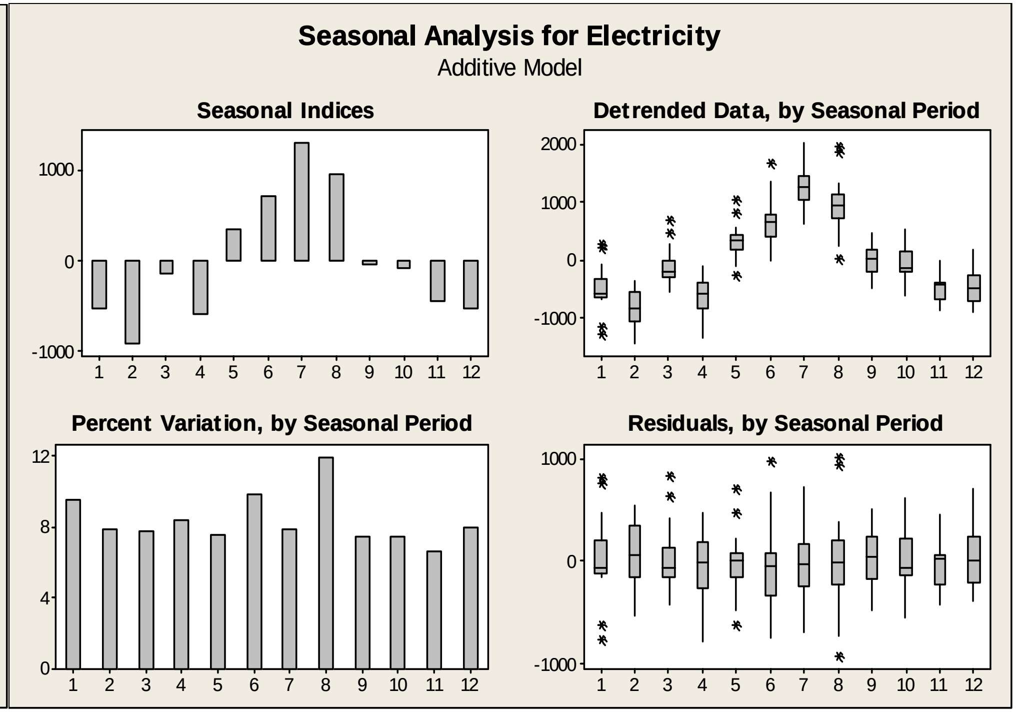
Comments

The deseasonlized data show some non-constant trend.
The residuals obtained by additional differencing with lag 1 of the differenced data with lag 12 are scattered about zero.
However, there are some bursts, which increase the variance of the residuals; the variance may not be constant.
The residuals may not be stationary.

MINITAB Decomposition Method







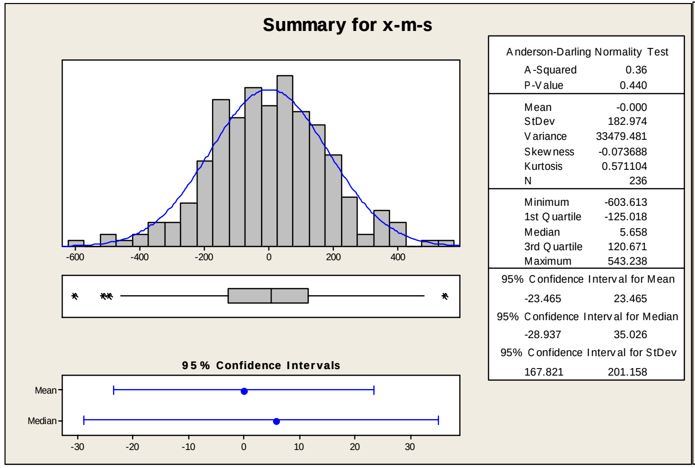
This method uses a linear trend approximation which, for this data set, does not completely remove trend from the data (as seen in the “Detrended Data” plot).

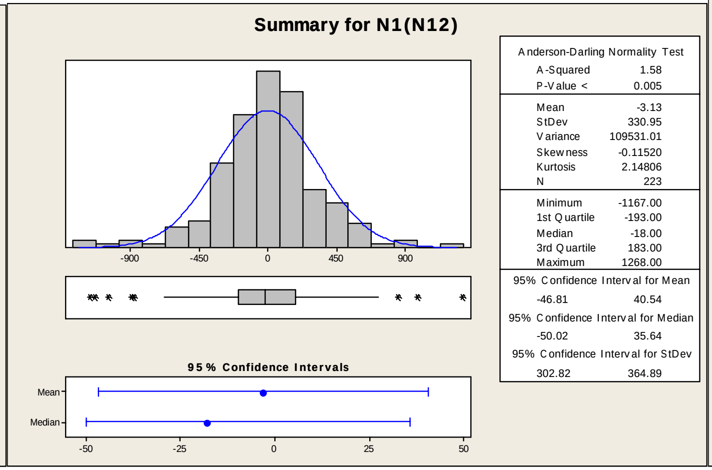
The deseasonlized data indicate an upward trend with a small bend around index 170 (year 1990)

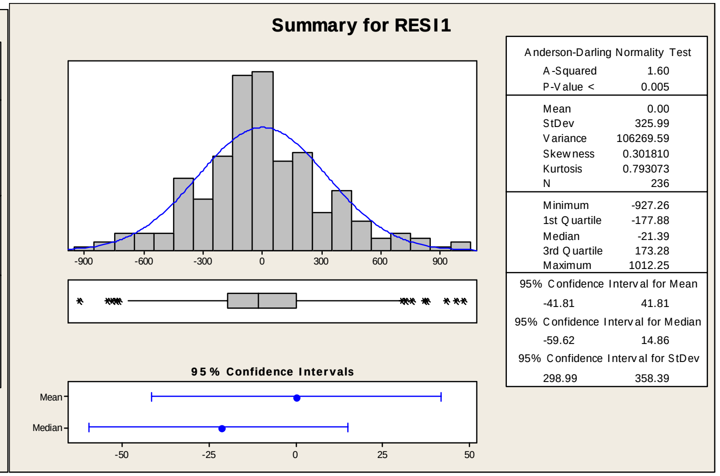
The residuals still carry the trend features.

Seasonal indices show the maximum production in July and minimum in February,
the highest seasonal variation in August,
although the variation by seasonal period is not large.

2.1.3 Comparison of the methods





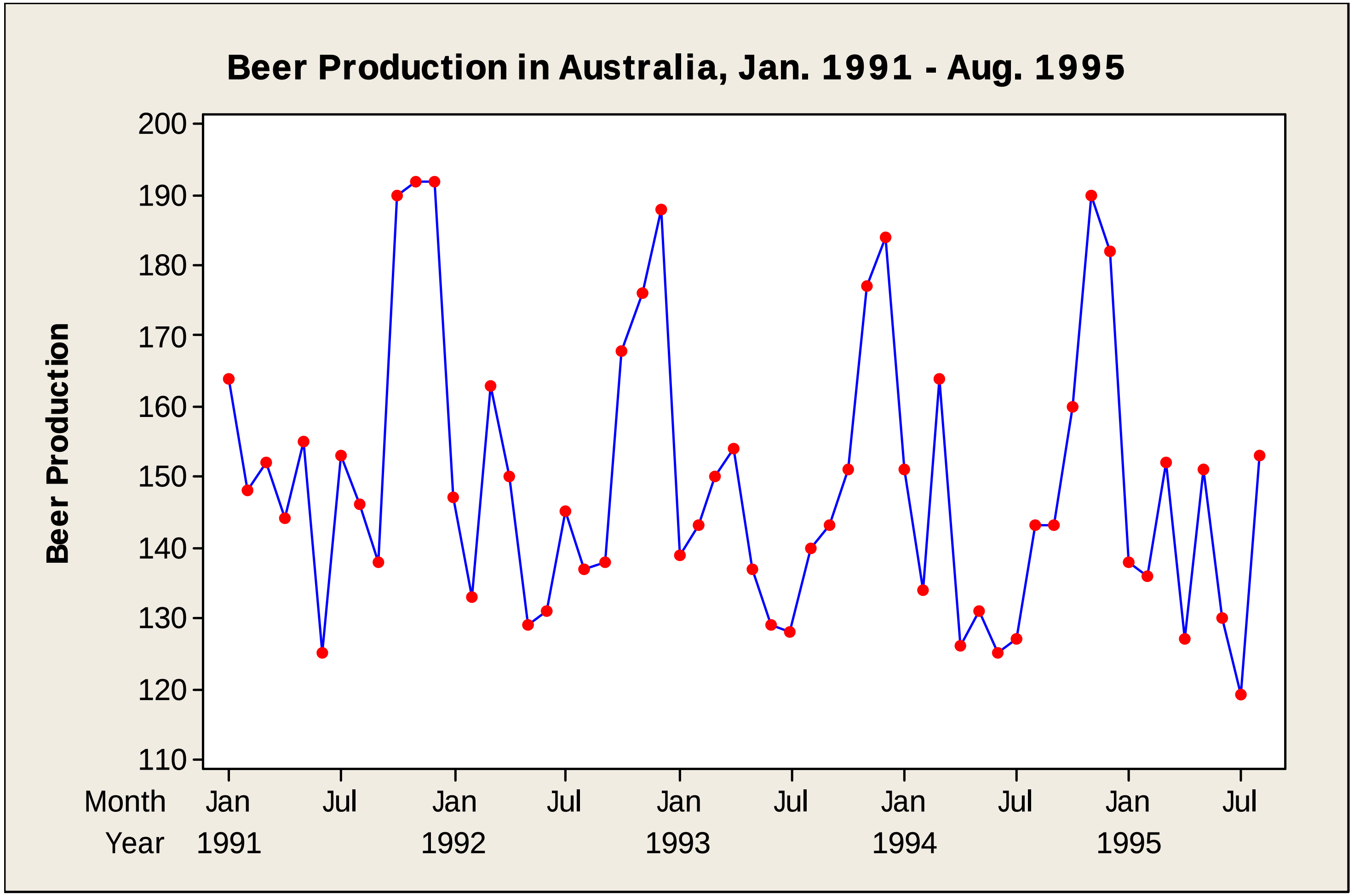


The histogram of the residuals for the Small Trend Method shows shape closest to a normal distribution.

The residuals have smallest variance and the test does not reject the hypothesis of normality.

Hence, this method seems to produce the best result, removed trend and seasonality most successfully.

2.1.4 Australian Beer Production

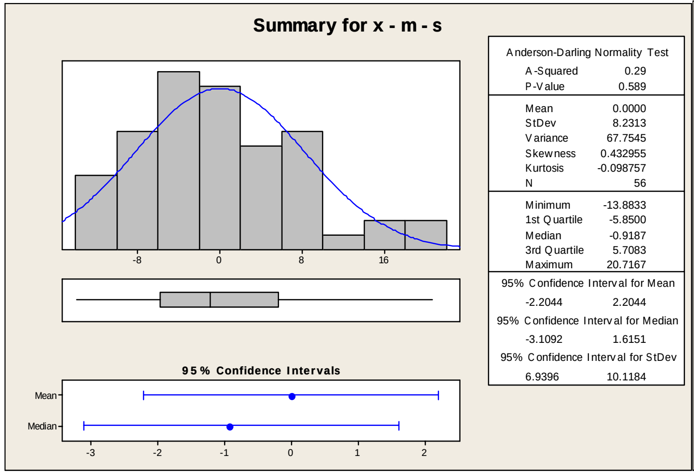


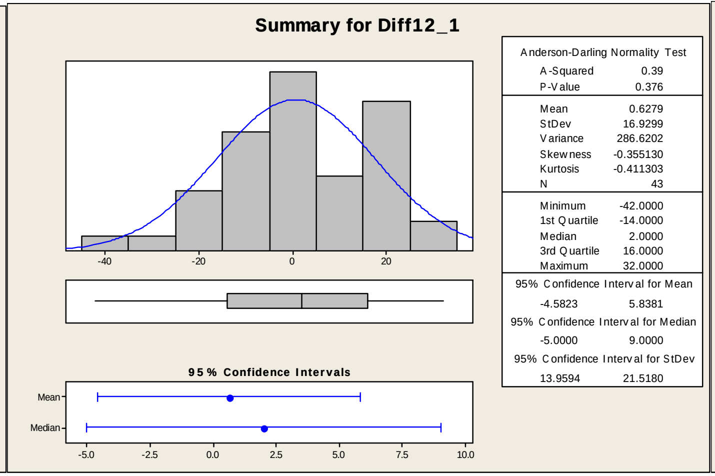
These are monthly observations, period length d = 12.

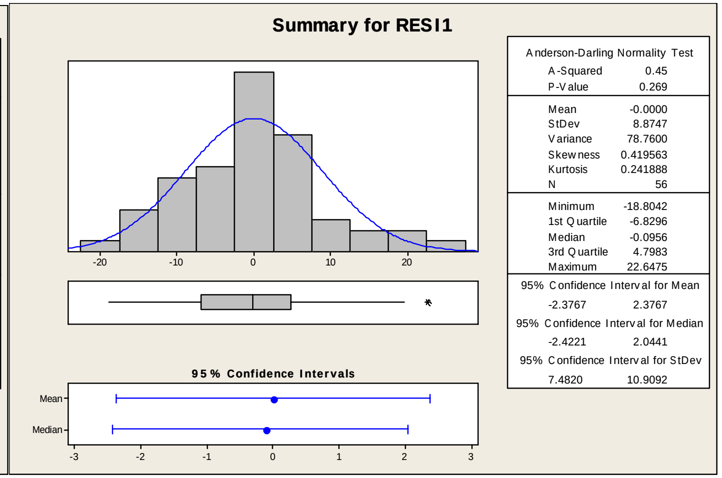
This data set shows possible seasonality effects and a small decreasing trend.

The beer production is highest in November and December, the summer months in Australia.

Comparison of the methods







For the Australian beer production data, there is no clear distinction among the methods, when compared via residuals.

The Small Trend Method gave the residuals of smallest range and largest p-value of the Anderson-Darling test of normality.