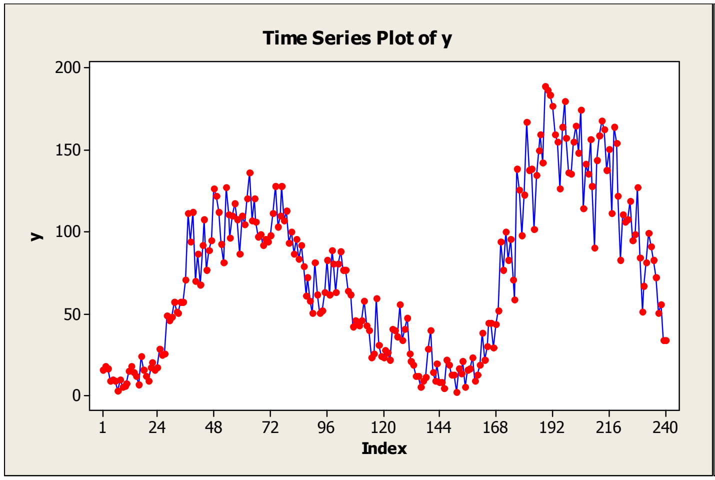
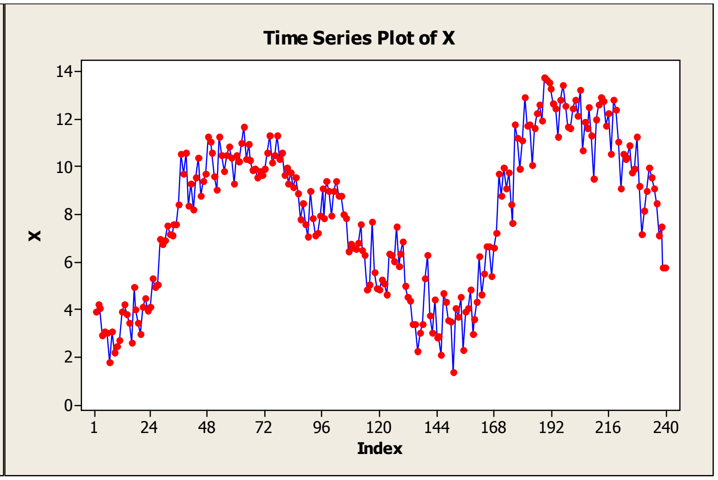
Minitab Project Report - Assignment 6

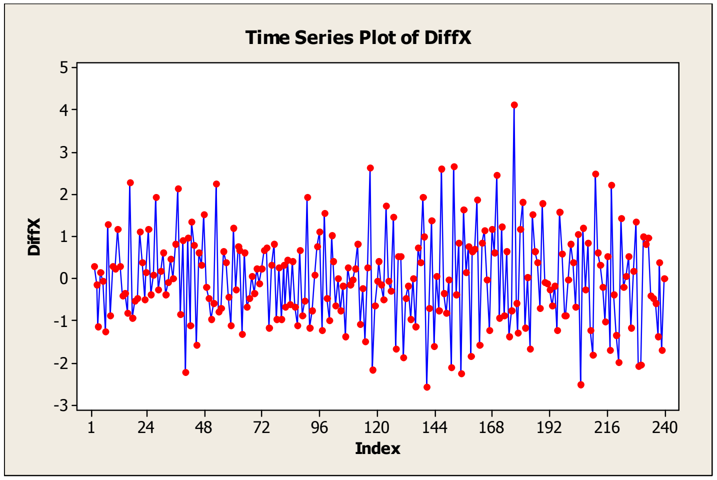
6.1.1 Sunspot data



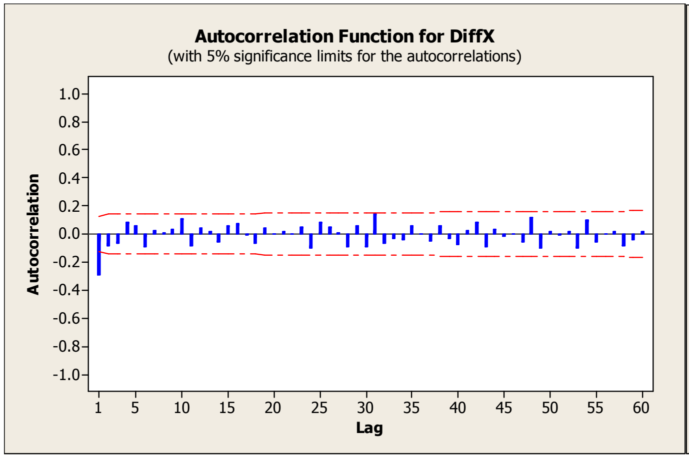


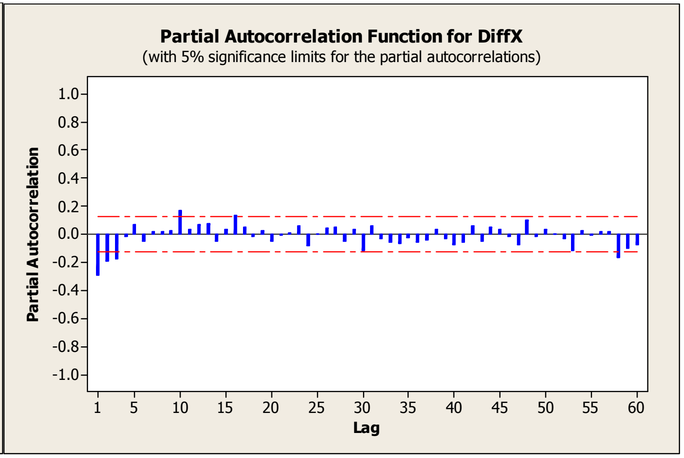
The data have a wavy pattern. However, they do not show any seasonality. There seem to be an increasing variability at the higher levels of the observations.

A power transformation (Box-Cox transformation performed in Minitab) stabilizes the variance.



DiffX: the differenced data represent the monthly changes in the transformed (by square root) numbers of sunspots.
DiffX series looks stationary with zero mean.





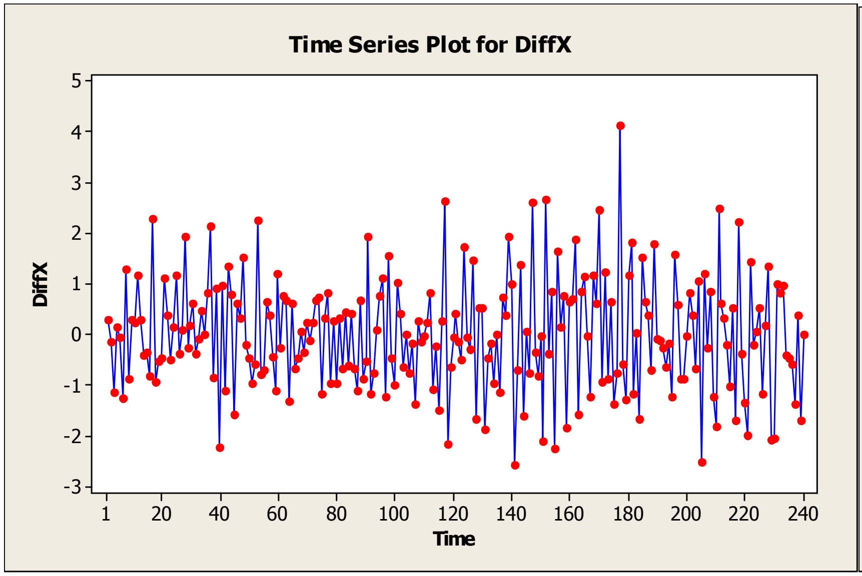
Autocorrelation Function: DiffX

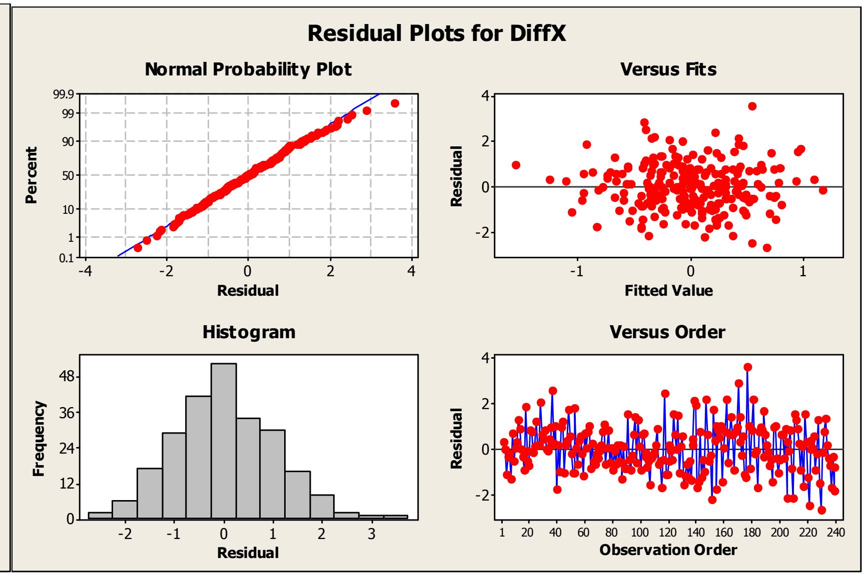
Lag	ACF	T	LBQ
1	-0.294591	-4.55	21.00
2	-0.086777	-1.24	22.83
3	-0.069790	-0.99	24.02
4	0.086939	1.23	25.87
5	0.062427	0.88	26.83
6	-0.095245	-1.33	29.08

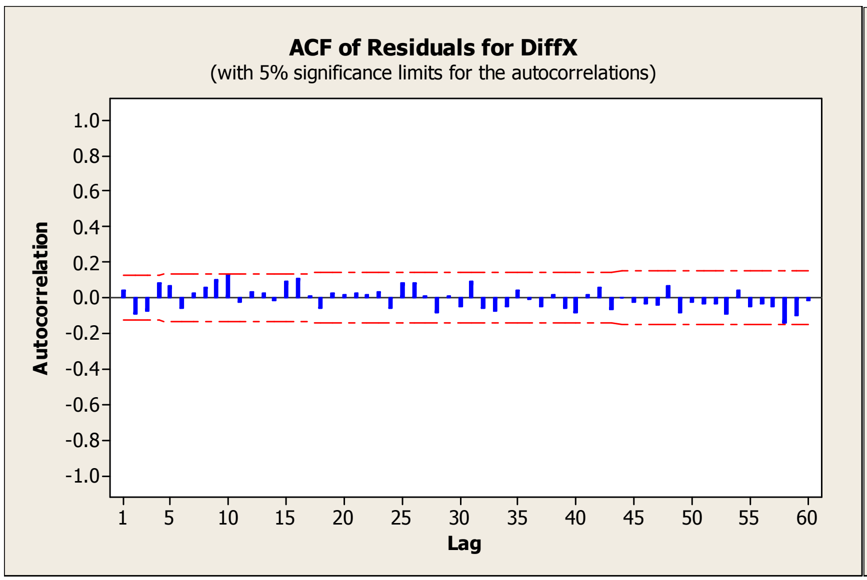
Partial Autocorrelation Function: DiffX

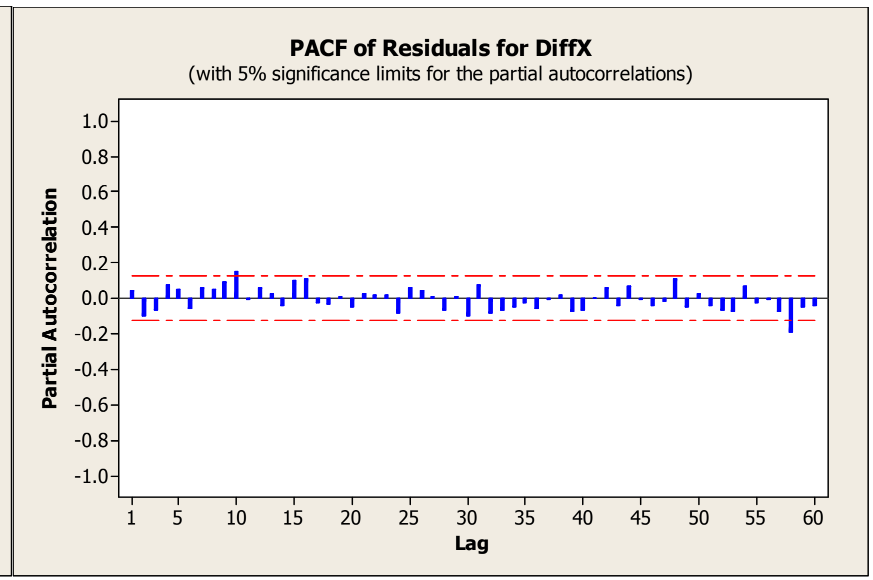
Lag	PACF	T
1	-0.294591	-4.55
2	-0.190054	-2.94
3	-0.177453	-2.74
4	-0.016479	-0.25
5	0.067342	1.04
6	-0.048177	-0.74

The sample ACF and PACF suggest an MA(1) model because the ACF cuts off at lag 1 and the PACF tails off. Also, $\hat{\rho}(1)$ is negative and all the significant sample PACF values are negative – this supports the choice of MA(1) with a negative estimate of θ.


Time Series Plot for DiffX
DiffX
Time

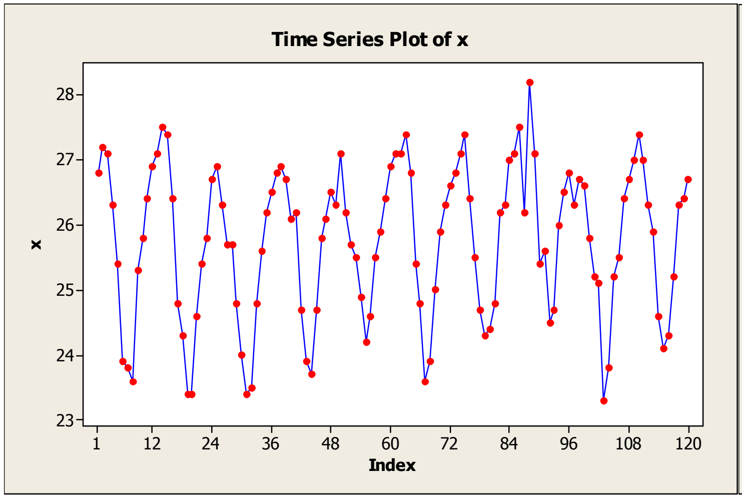

Residual Plots for DiffX
Normal Probability Plot
Percent
Residual
Versus Fits
Residual
Fitted Value
Histogram
Frequency
Residual
Versus Order
Residual
Observation Order

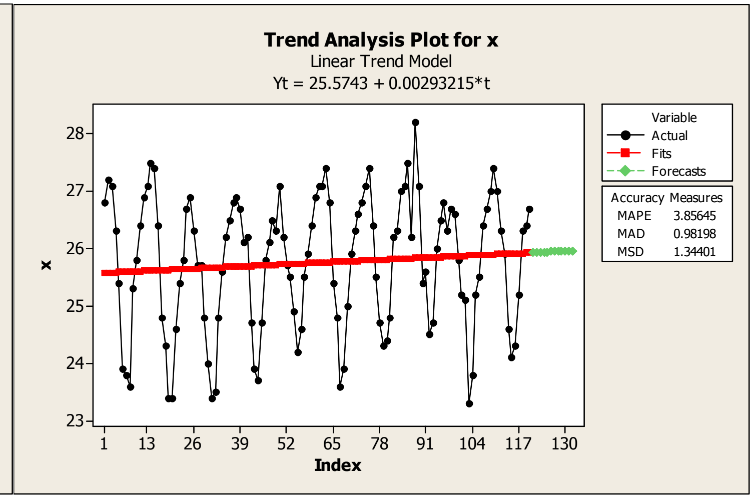

ACF of Residuals for DiffX
(with 5% significance limits for the autocorrelations)
Autocorrelation
Lag


PACF of Residuals for DiffX
(with 5% significance limits for the partial autocorrelations)
Partial Autocorrelation
Lag

Modelling MA(1)
 	
`Type      Coef  SE Coef     T      P`
`MA   1  0.4341   0.0588  7.38  0.000`
 	
Modified Box-Pierce (Ljung-Box) Chi-Square statistic
`Lag             12     24     36     48`
`Chi-Square    16.3   24.2   36.6   45.2`
`DF              11     23     35     47`
`P-Value      0.132  0.390  0.395  0.54`
	The model parameter θ is statistically significant (p<0.000), $\hat{\theta}$ = - 0.4341 (note that Minitab is using Box and Jenkins' (1976) notation for MA(1), that is $X_t = Z_t - \theta Z_{t-1}$).
	
	Hence, the MA(1) model is $\tilde{X}_t = Z_t - 0.4341\, Z_{t-1}$, where $\tilde{X}_t$ denotes the transformed and differenced series, Z_t denotes a white noise.
	
	The residuals look like Gaussian White Noise.
	Also, the Box-Pierce (Ljung-Box) Chi-Square statistics show that the residuals of this model (in groups of up to 48 values) are uncorrelated.
	
	MA(1) fits well the transformed and differenced sunspots data.
	
	Another model, such as AR(3) might also be considered.

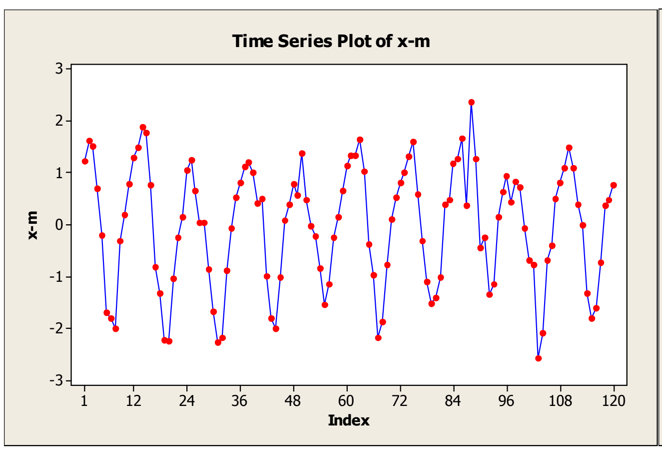
6.1.2 Recife Data

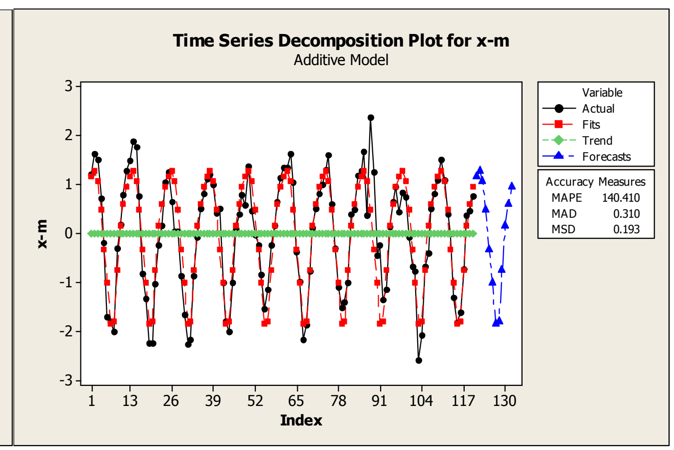




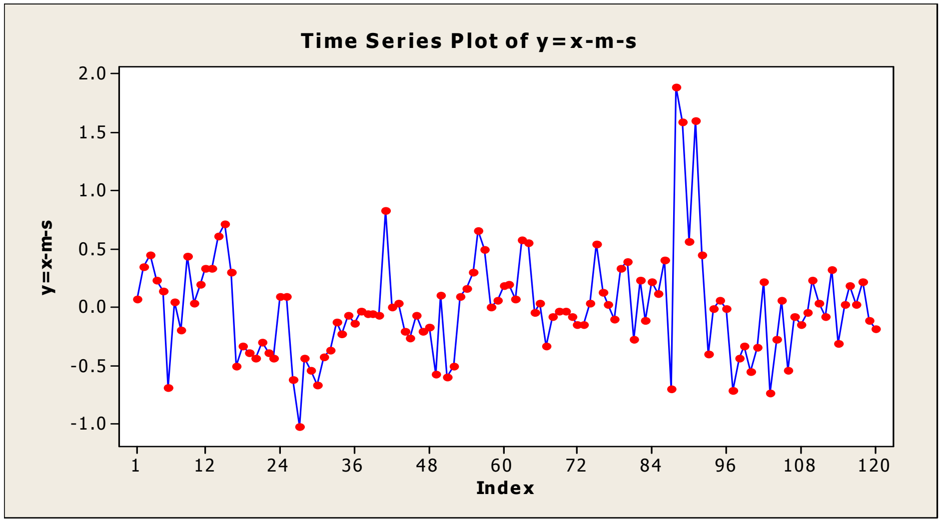
Monthly average temperature in Recife in Celsius in years 1953-62. The temperature varies between about 23 – 28 degrees C.
The warmest months are December-March and least warm are June-August.
There is clear seasonality in the data and not very clear, perhaps slightly upward, trend.
Unusual pattern occurred in the eighth year of the recorded data.

The linear trend fit shows a small increase in temperature over the years.

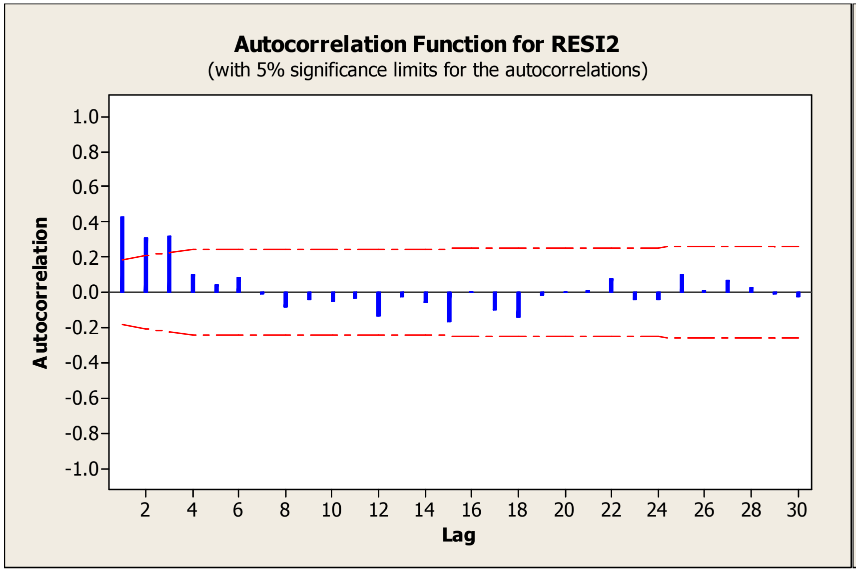


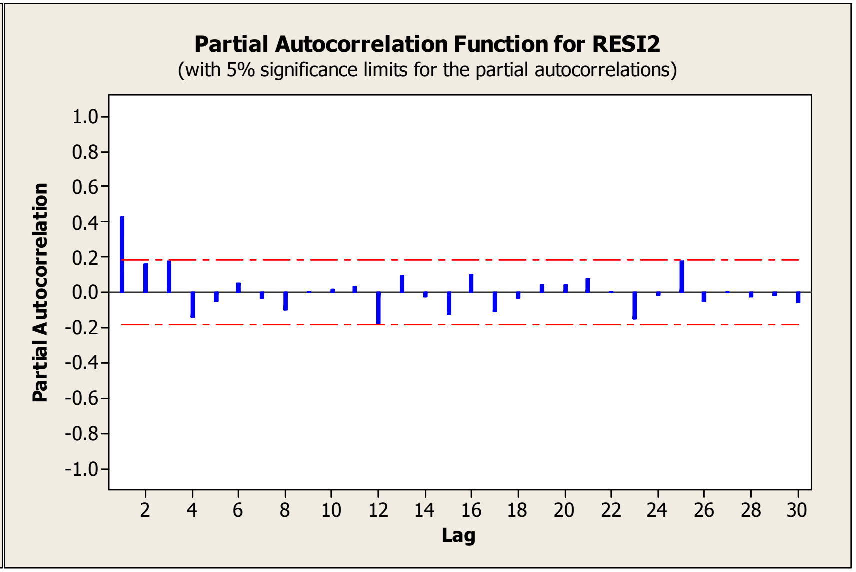


The detrended data exhibit seasonality but no trend, the 12 forecasted values repeat the pattern of the seasonal effects.



The seasonally adjusted and detrended data seem to be correlated with a few unusual values reflecting the pattern in the eighth year of the recorded data.

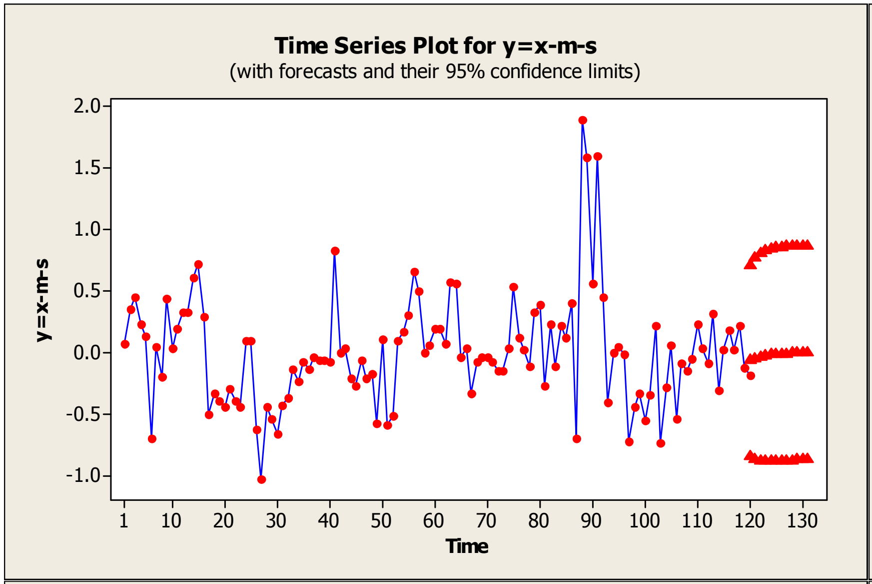


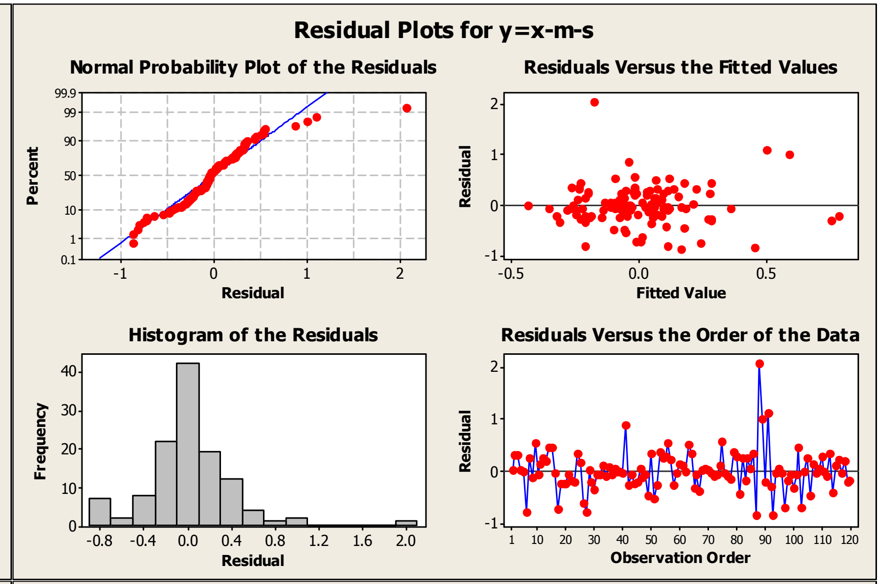


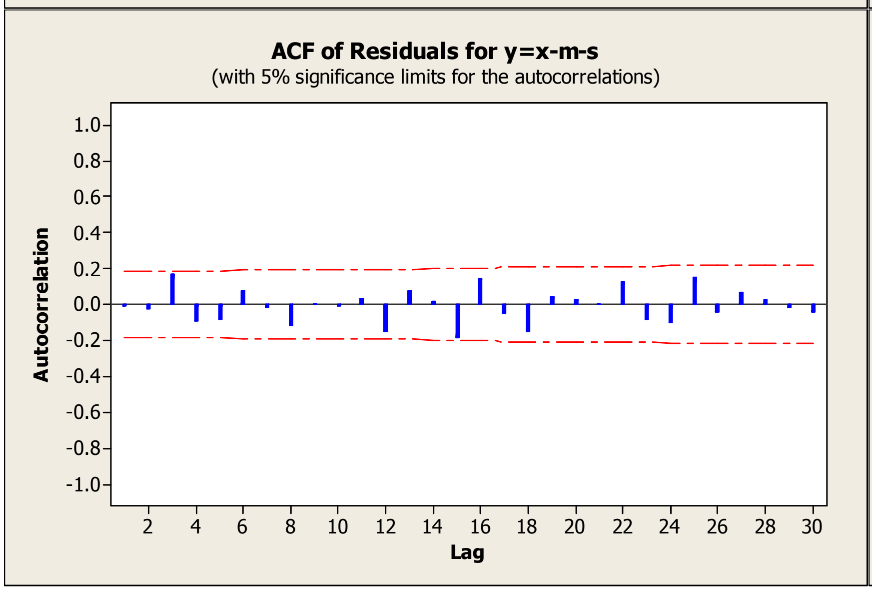
The sample ACF and PACF suggest an ARMA process. Trying various models we have obtained a reasonable fit with ARMA(1,1):

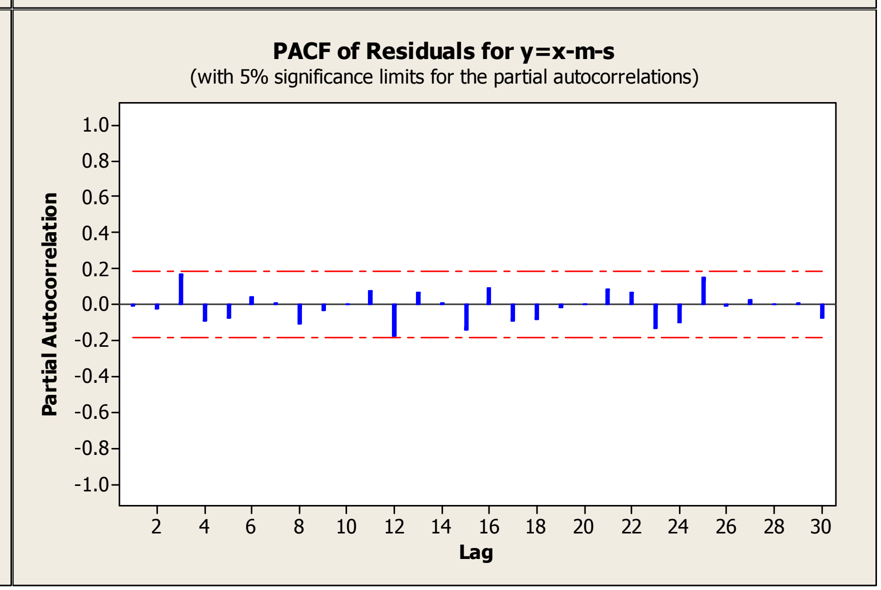
$$X_t - \varphi X_{t-1} = Z_t + \vartheta Z_{t-1}, \quad Z_t \sim WN(0, \sigma^2).$$

Below are given plots of the detrended and deseasonalised data and their 12 forecasted values. Also, there are diagnostic plots for checking the white noise assumptions of the ARMA(1,1) model. As we can see the assumptions are approximately met.

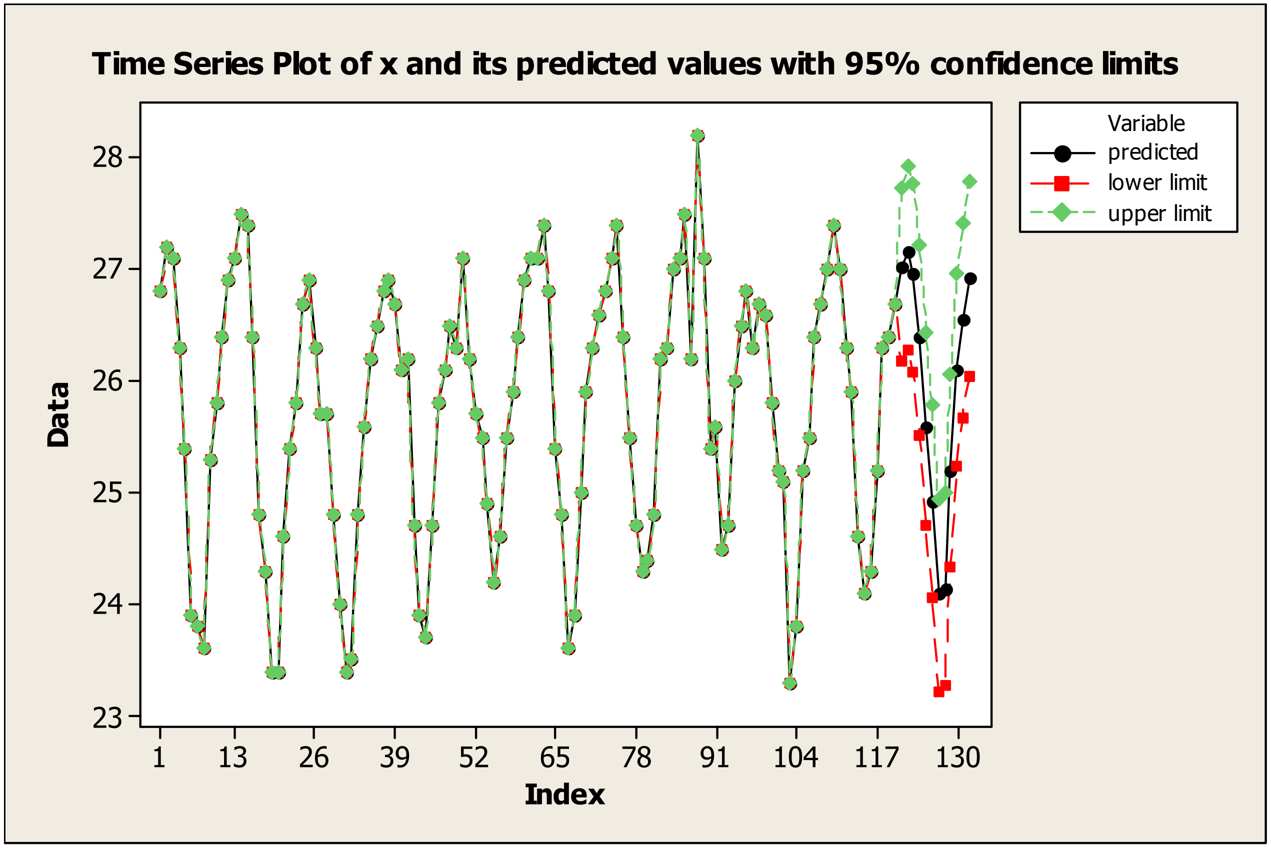








Here is the plot of the original data and the 12 added forecasted values.



The numerical Minitab output:

Trend Analysis for x

Fitted Trend Equation
Yt = 25.5743 + 0.00293215*t

Period	Forecast
121	25.9291
122	25.9320
123	25.9349
124	25.9379
125	25.9408
126	25.9437
127	25.9467
128	25.9496
129	25.9525
130	25.9555
131	25.9584
132	25.9613

Time Series Decomposition for x-m

Additive Model

Period	Forecast
121	1.15694
122	1.27361
123	1.06944
124	0.48194
125	-0.32222
126	-0.99722
127	-1.83889
128	-1.79722
129	-0.73889
130	0.15694
131	0.59444
132	0.96111

ARIMA Model: y=x-m-s

Type		Coef	SE Coef	T	P
AR	1	0.7378	0.1278	5.77	0.000
MA	1	0.3911	0.1743	2.24	0.027

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

Lag	12	24	36	48
Chi-Square	11.1	28.2	35.9	56.5
DF	10	22	34	46
P-Value	0.353	0.171	0.378	0.139

Forecasts from period 120

		95 Percent Limits		
Period	Forecast	Lower	Upper	Actual
121	-0.067285	-0.840540	0.705970	
122	-0.049643	-0.868060	0.768775	
123	-0.036626	-0.878609	0.805357	
124	-0.027022	-0.881561	0.827516	
125	-0.019937	-0.881232	0.841358	
126	-0.014709	-0.879661	0.850242	
127	-0.010853	-0.877788	0.856083	
128	-0.008007	-0.876020	0.860006	
129	-0.005907	-0.874507	0.862692	
130	-0.004358	-0.873277	0.864560	
131	-0.003216	-0.872307	0.865876	
132	-0.002372	-0.871559	0.866814	