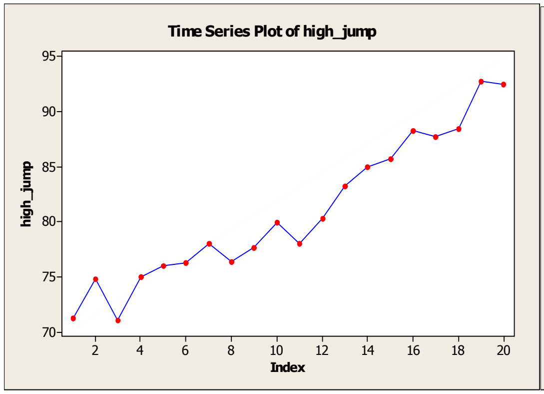
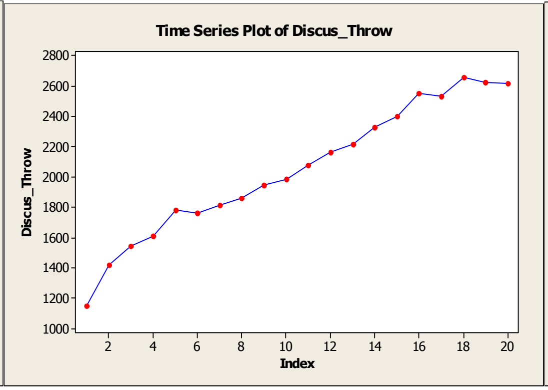
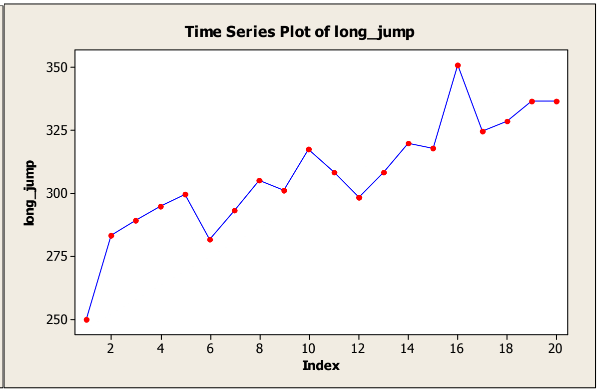
1.3 Time series plot







All the data for the three Olympic events have **no seasonality** by their nature, however the plots indicate strong **increasing trend** in each series.

None of the series has a noise component increasing in time.

There are no sudden peaks or clear turning points.

The Discus Throw is most 'smooth', while High Jump and Long Jump results, although increasing, have jumps up and down.

1.4 Trend analysis
1.4.1 Regression model fit

Accuracy measures

MAPE Mean Absolute Percentage Error (MAPE) measures the accuracy of fitted time series values. It expresses accuracy as a percentage.

$$\text{MAPE} = \frac{1}{n}\sum\left|(y_t - \hat{y}_t)/y_t\right|100$$

where y_t equals the actual value at time t, $\hat{y}_t$ equals the fitted value, and n equals the number of observations.

MAD Mean Absolute Deviation (MAD) measures the accuracy of fitted time series values. It expresses accuracy in the same units as the data, which helps conceptualize the amount of error.

$$\text{MAD} = \frac{1}{n}\sum\left|y_t - \hat{y}_t\right|$$

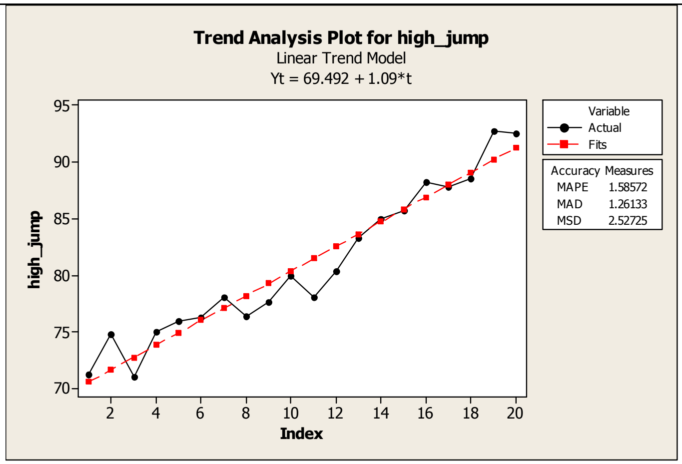
where y_t equals the actual value at time t, $\hat{y}_t$ equals the fitted value, and n equals the number of observations.

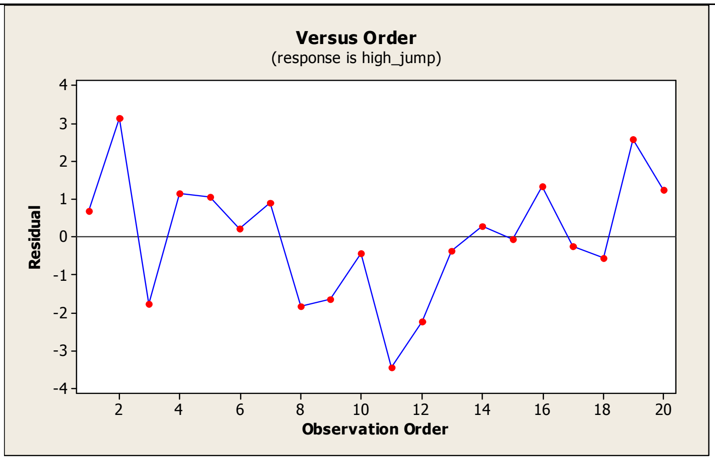
MSD Mean Squared Deviation (MSD) is always computed using the same denominator, n, regardless of the model. MSD is a more sensitive measure of an unusually large forecast error than MAD.

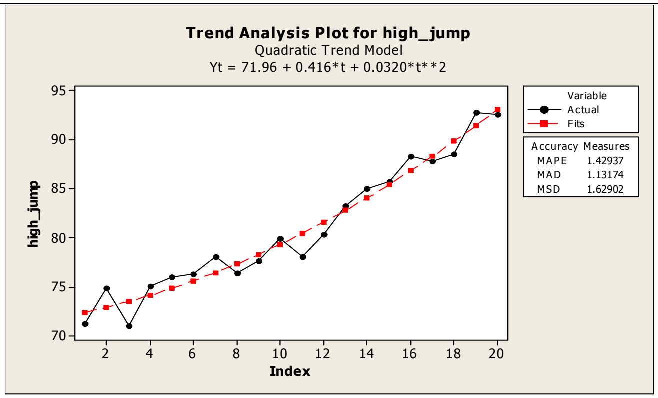
$$\text{MSD} = \frac{1}{n}\sum(y_t - \hat{y}_t)^2$$

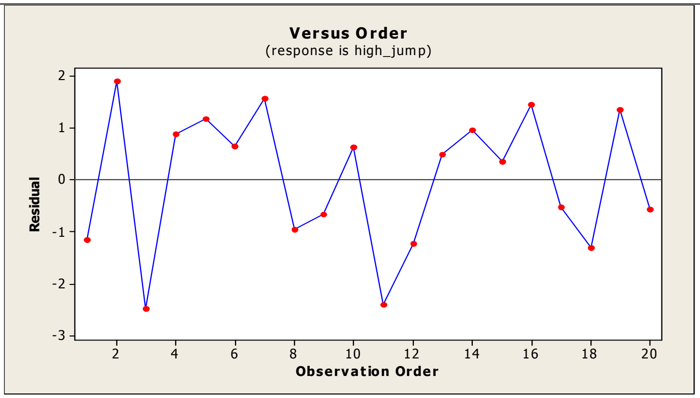
where y_t equals the actual value at time t, $\hat{y}_t$ equals the forecast value, and n equals the number of forecasts.

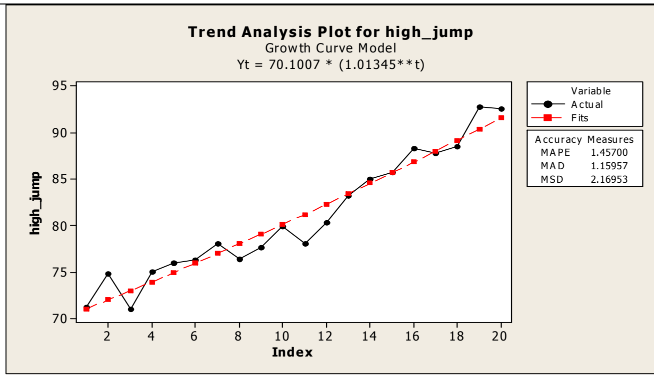
Trend Analysis for high_jump

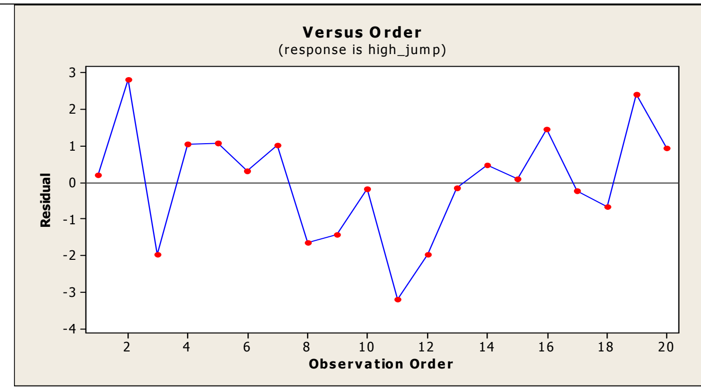


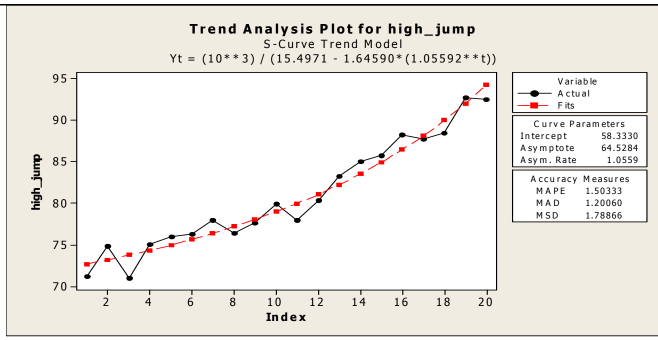


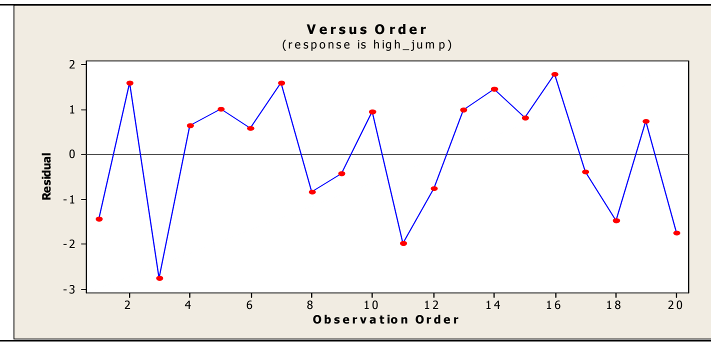












Fitted Trend Equation Yt = 69.492+1.09*t Accuracy Measures MAPE 1.58572 MAD 1.26133 MSD 2.52725	Fitted Trend Equation Yt = 71.96+0.416*t + 0.0320*t**2 Accuracy Measures MAPE **1.42937** MAD **1.13174** MSD **1.62902**	Fitted Trend Equation Yt = 70.1007* (1.01345**t) Accuracy Measures MAPE 1.45700 MAD 1.15957 MSD 2.16953	Fitted Trend Equation Yt =(10**3)/(15.4971-1.64590*(1.05592**t)) Accuracy Measures MAPE 1.50333 MAD 1.20060 MSD 1.78866

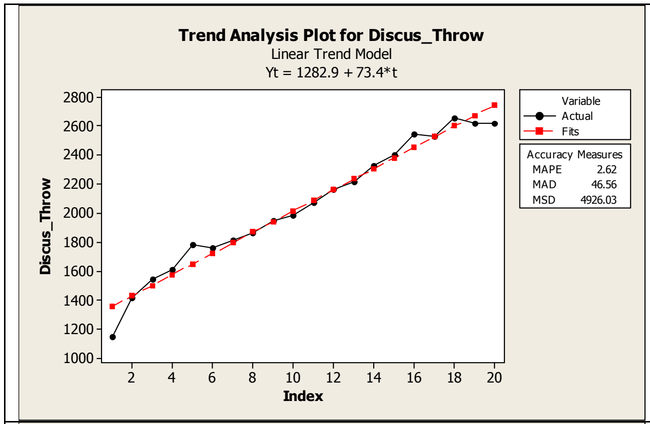
Comments:

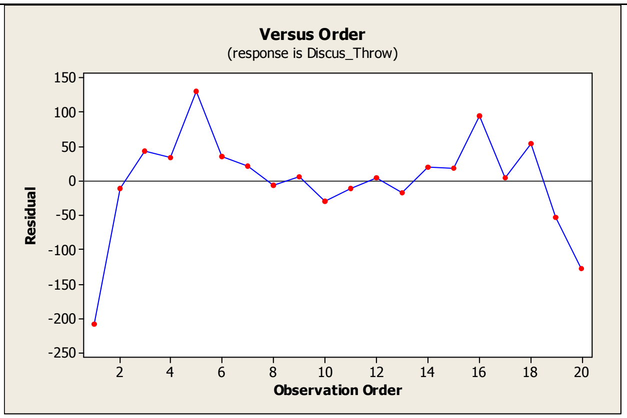
Model chosen: $X_t = \beta_0 + \beta_1 t + \beta_2 t^2 + Y_t$ where Y_t denotes random noise.

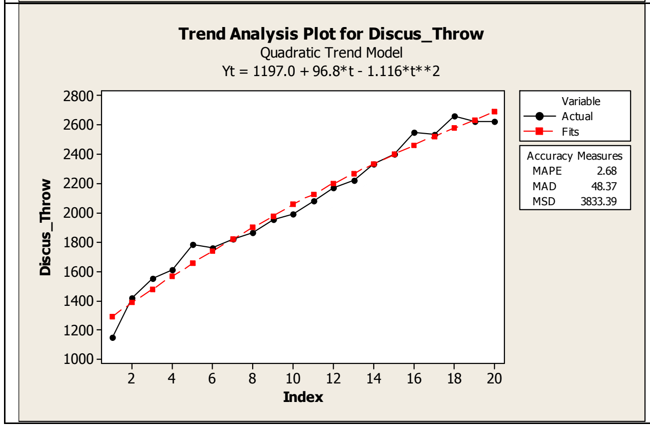
The smallest values of the model accuracy measures are for the quadratic trend, the residuals seem to be scattered about zero with no indication of trend or of non-constant variance.

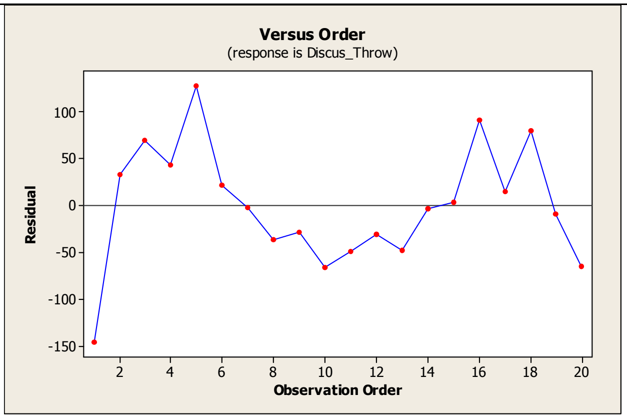
None of the models might be good for prediction of far future values of the high jump, as they all sharply increase, what may not be realistic, though they may be fine for prediction of the value for next Olympic game. We would expect the increase in the performance to 'slow down' simply due to limitations of human body.

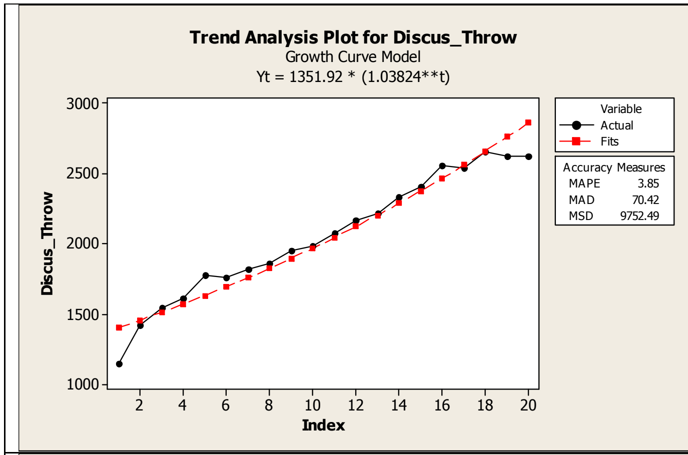
Trend Analysis for Discus_Throw

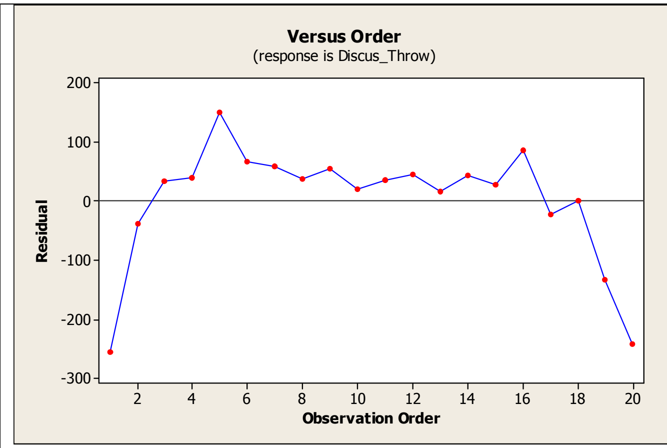


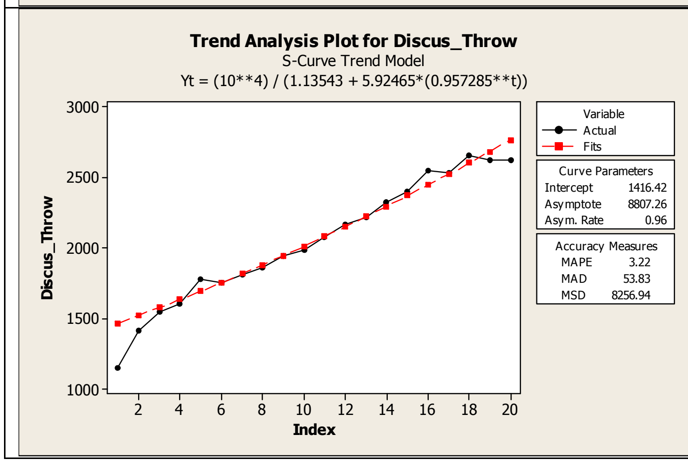


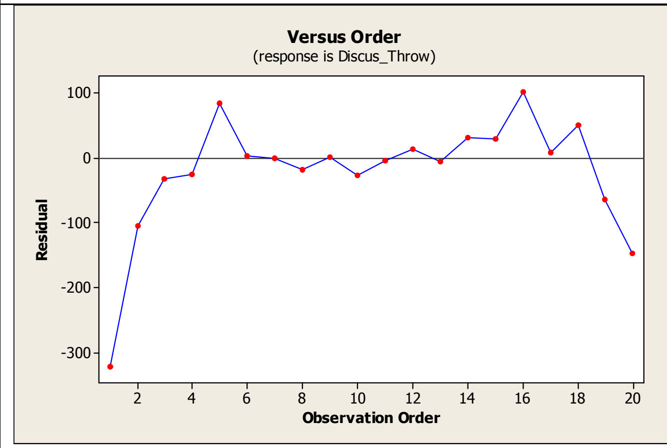












Fitted Trend Equation Yt = 1282.9 + 73.4*t Accuracy Measures MAPE 2.62 MAD 46.56 MSD 4926.03	Fitted Trend Equation Yt = 1197.0 + 96.8*t - 1.116*t**2 Accuracy Measures MAPE 2.68 MAD 48.37 MSD 3833.39	Fitted Trend Equation Yt = 1351.92 * (1.03824**t) Accuracy Measures MAPE 3.85 MAD 70.42 MSD 9752.49	Fitted Trend Equation Yt = (10**4) / (1.13543 + 5.92465*(0.957285**t)) Accuracy Measures MAPE 3.22 MAD 53.83 MSD 8256.94

Commetents

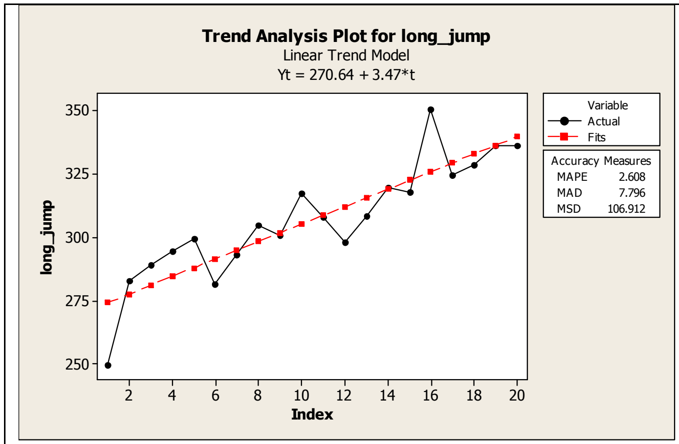
Model chosen: $X_t = \beta_0 + \beta_1 t + \beta_2 t^2 + Y_t$ where Y_t denotes random noise.

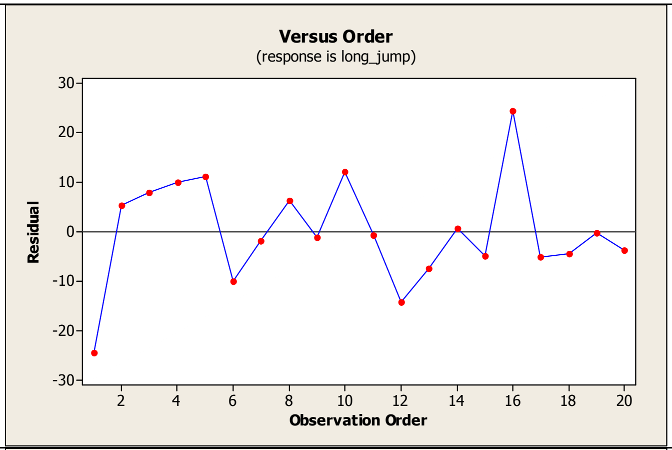
None of the trend models has all three accuracy measures best. Linear trend has smallest MAPE and MAD, while the quadratic trend has smallest MSD. The other two trend models have larger measures, particularly MSD.

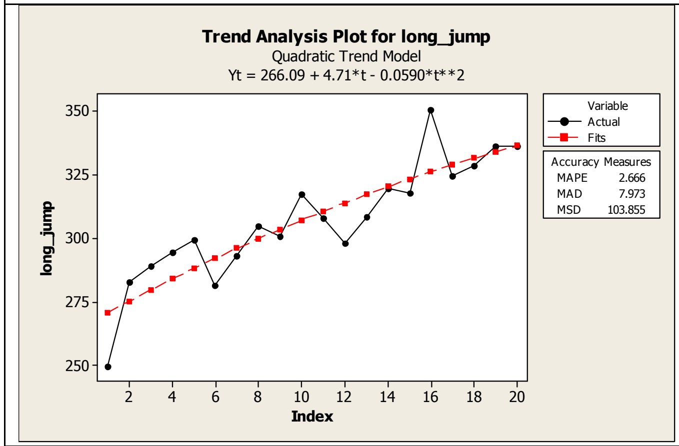
Comparing the linear and quadratic trends, we see that

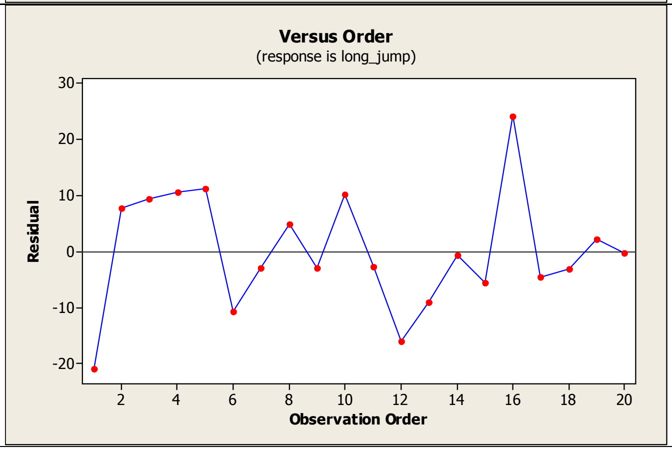
- there is not much difference between MAPE and MAD criteria,
- MSD is much lower for the quadratic model,
- the quadratic trend fits better at the ends of the series, particularly at the last value. It may give better prediction values for near future,
- in general, its residuals are less varied.

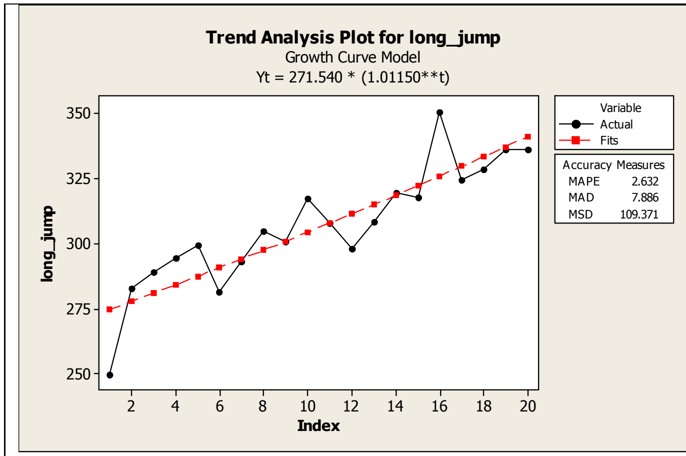
Trend Analysis Plot for long_jump

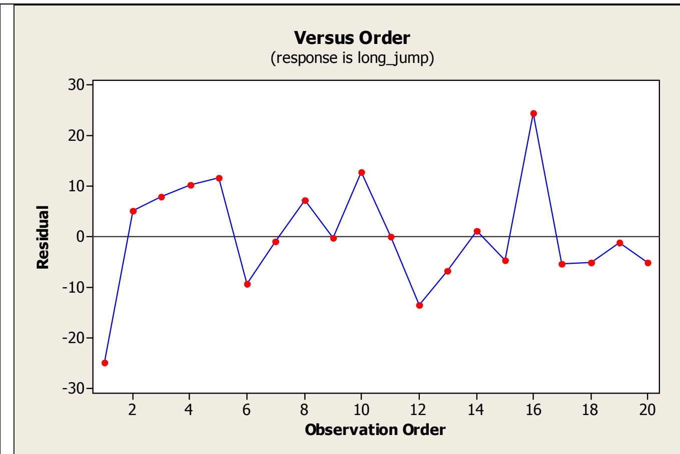


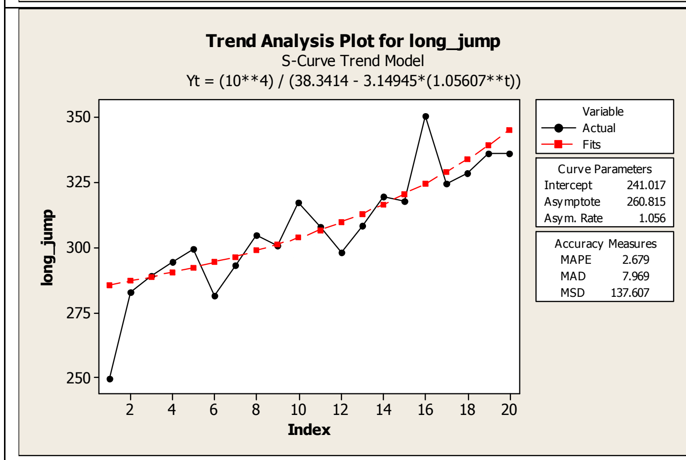


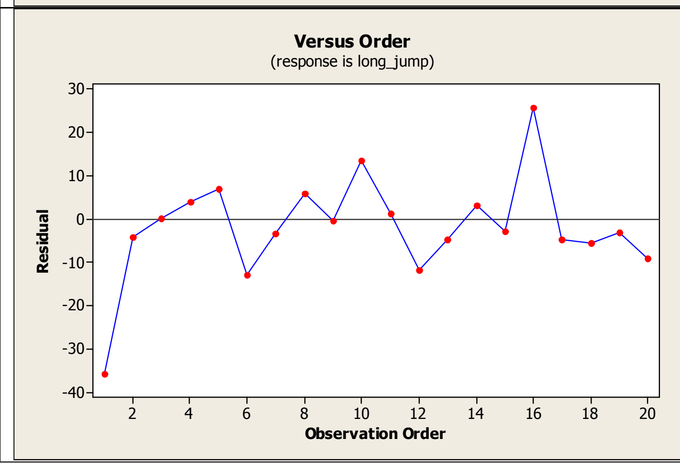












Fitted Trend Equation Yt = 270.64 + 3.47*t Accuracy Measures MAPE 2.608 MAD 7.796 MSD 106.912	Fitted Trend Equation Yt = 266.09 + 4.71*t - 0.0590*t**2 Accuracy Measures MAPE 2.666 MAD 7.973 MSD 103.855	Fitted Trend Equation Yt = 271.540 * (1.01150**t) Accuracy Measures MAPE 2.632 MAD 7.886 MSD 109.371	Fitted Trend Equation Yt = (10**4) / (38.3414 - 3.14945*(1.05607**t)) Accuracy Measures MAPE 2.679 MAD 7.969 MSD 137.60

Comments

Model chosen: $X_t = \beta_0 + \beta_1 t + \beta_2 t^2 + Y_t$ where Y_t denotes random noise.

None of the trend models has all three accuracy measures best. Linear trend has smallest MAPE and MAD, while the quadratic trend has smallest MSD. The other two trend models have larger values of the measures.

Comparing the linear and quadratic trends, we see that

- there is not much difference between the criteria,
- the quadratic trend fits better at the end of the series, particularly at the last value. It may give better prediction values for near future,
- in general, its residuals of all trend models look stationary, but the quadratic model gives residuals on a smaller range.

Trend models (careful with notation)

Linear — Trend analysis by default uses the *linear trend* model:

$Y_t = \beta_0 + \beta_1 t + e_t$

In this model, β_1 represents the average change from one period to the next.

Quadratic — The *quadratic trend model* which can account for simple curvature in the data, is:

$Y_t = \beta_0 + \beta_1 * t + \beta_2 t^2 + e_t$

Exponential Growth — The *exponential growth trend model* accounts for exponential growth or decay. For example, a savings account might exhibit exponential growth. The model is:

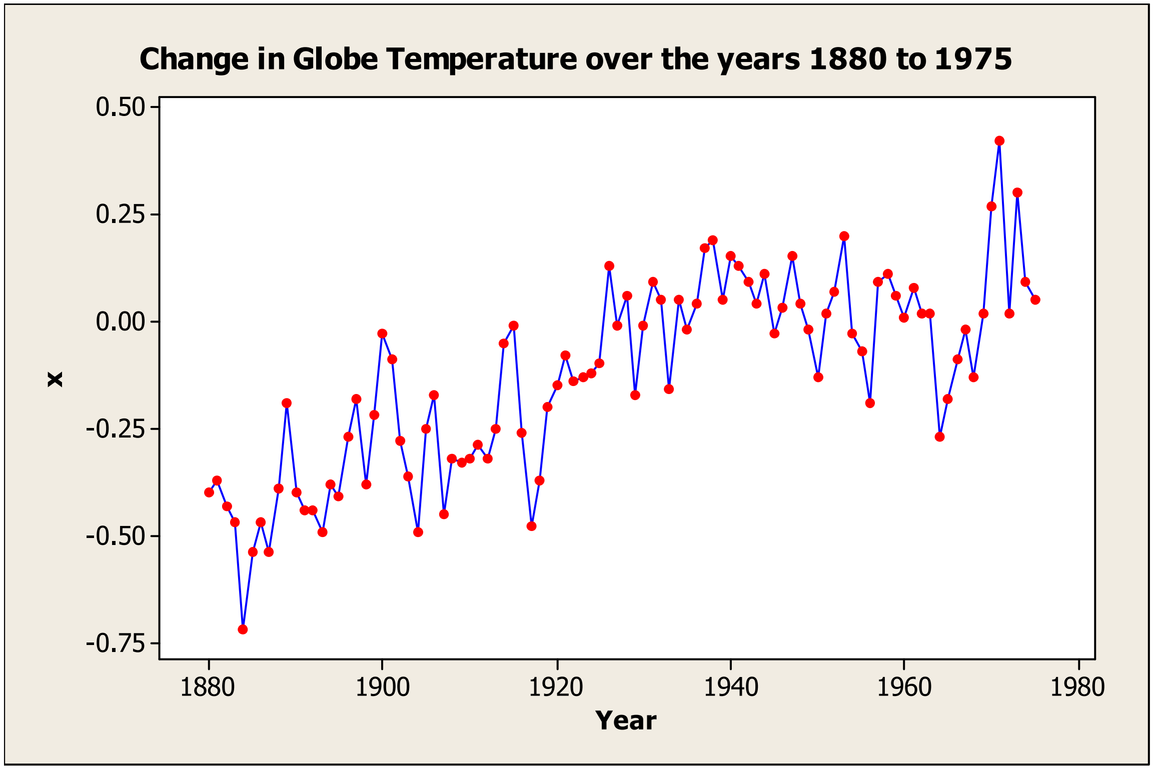
$Y_t = \beta_0 * \beta_1^t * e_t$

S-curve — The *S-curve model* fits the Pearl-Reed logistic trend model. This accounts for the case where the series follows an S-shaped curve. The model is:

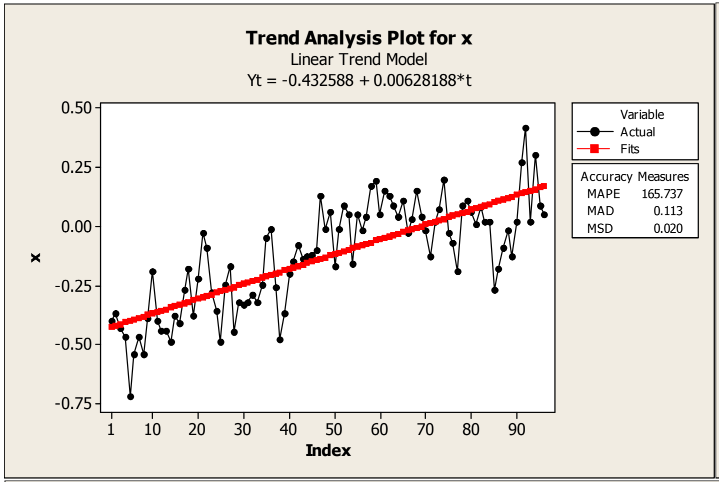
$Y_t = 10^a / (\beta_0 + \beta_1 \beta_2^t) + e_t$

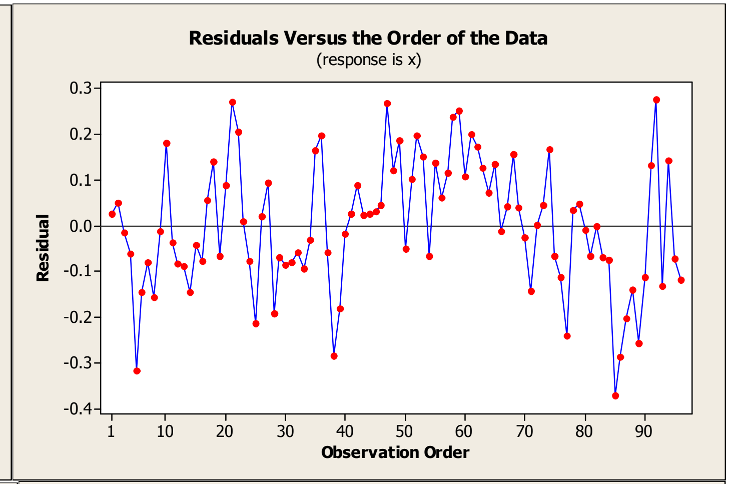
Note that the Exponential Growth model is multiplicative and so the residuals are the observations divided by the fitted values of the model. What we obtain as 'residuals' in Minitab is the difference between the two. The shape of the plot of the residuals is kept, but not the scale.

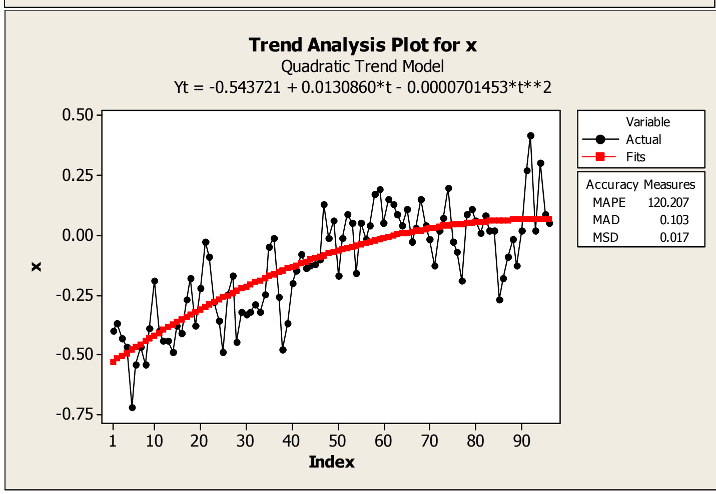
Change in Temperature Data

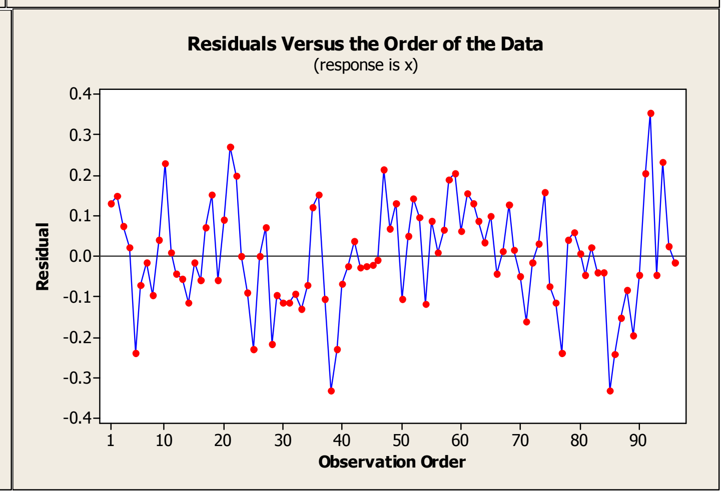


Comments:
There is steady upward trend in the change of temperature data.









Trend Analysis for x

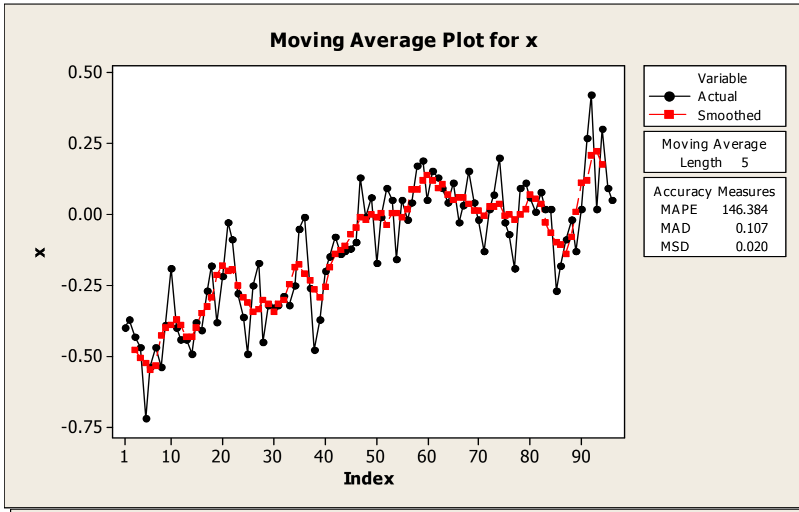
```
Yt = -0.432588 + 0.00628188*t
Accuracy Measures
MAPE  165.737
MAD     0.113
MSD     0.020

Yt = -0.543721 + 0.0130860*t - 0.0000701453*t**2
Accuracy Measures
MAPE  120.207
MAD     0.103
MSD     0.017
```

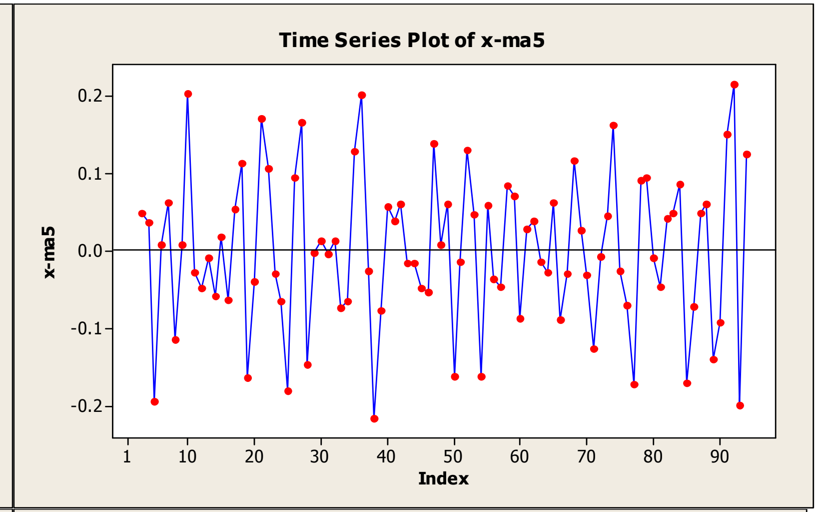
The accuracy measures give smaller values for the quadratic trend fit.

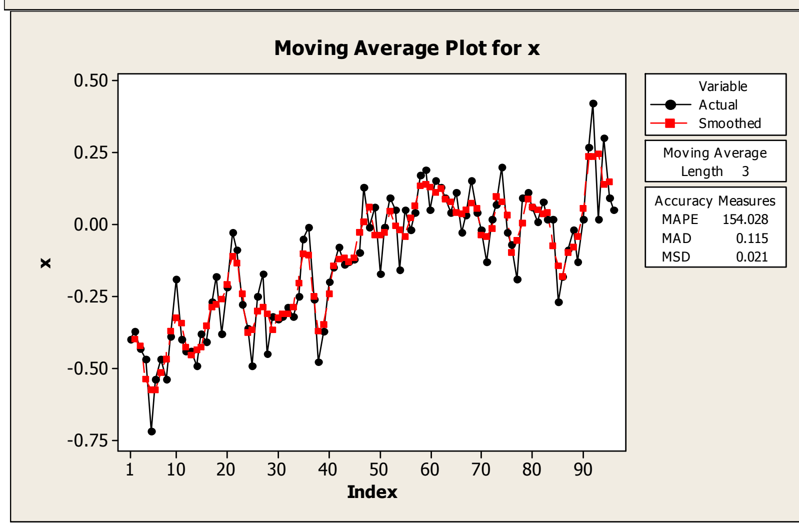
The growth curve and S-curve could not be fitted, as they assume that the values are all positive, while this TS has both negative and positive values.

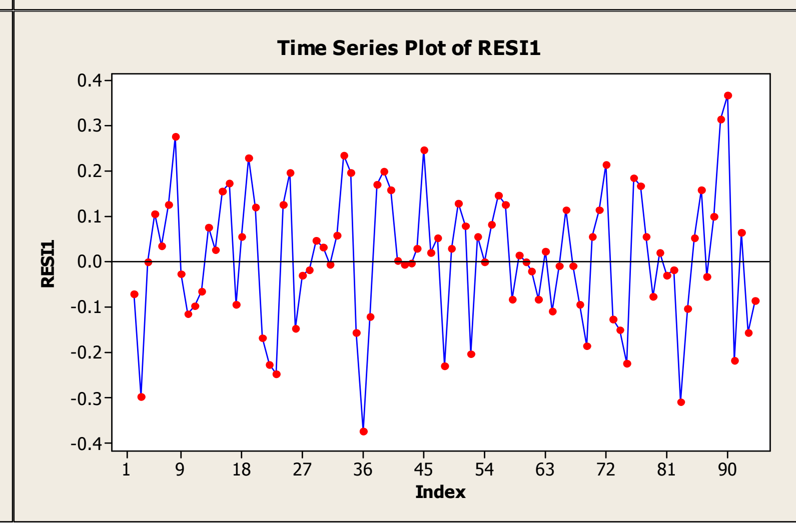
Indeed, the quadratic trend seems to fit quite well, the residuals show less wavy pattern than those for the linear trend.

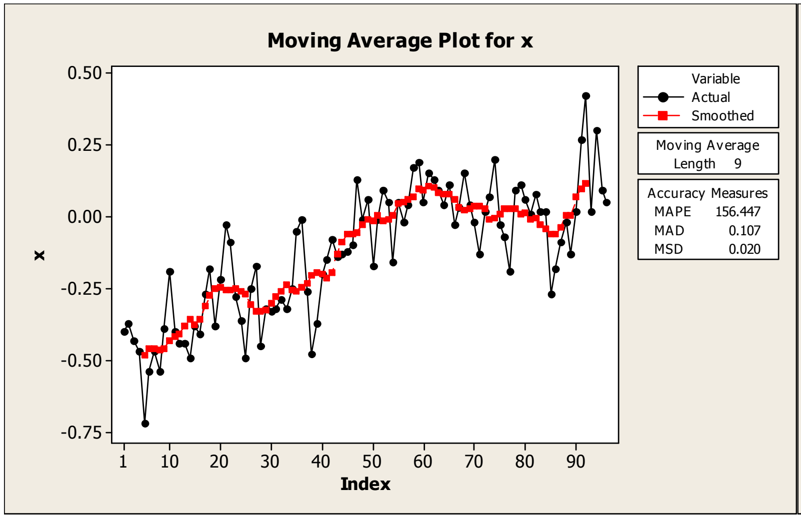
1.4.2 Moving Average Change in Temperature Data

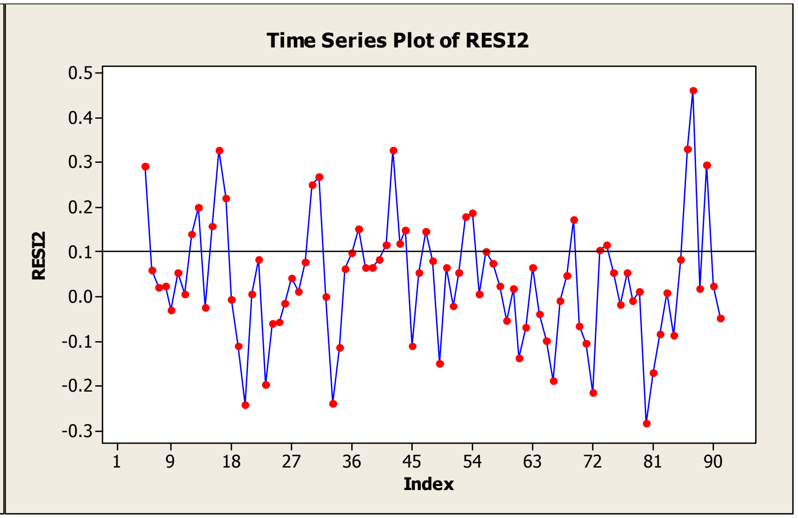












Obviously, the smaller is the length of m.a. the closer it is to the original series, and so the residuals have smaller range and are less bursty.

The m.a. of length 9 is the smoothest among the three but gives most bursty residuals, reflecting the burstiness of the original TS.