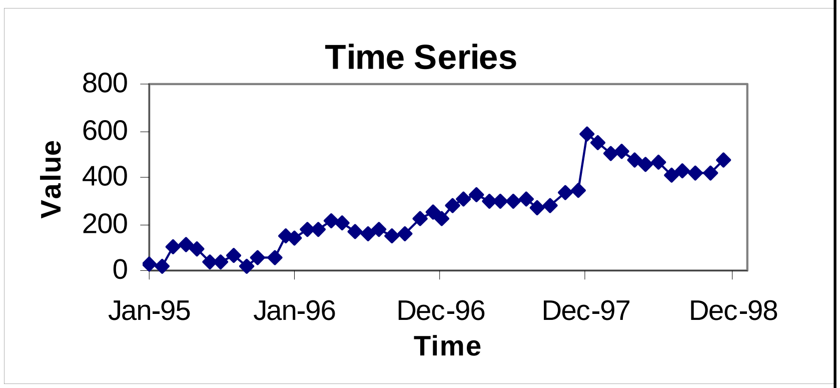
Module 8: Time Series Analysis

8.2 Components of a Time Series, Detection of Change Points and Trends, Time Series Models

Components of a Time Series

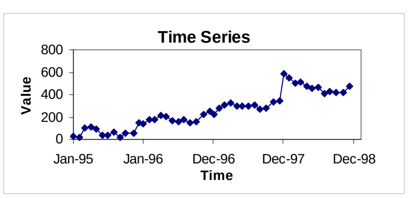
- There can be several things happening simultaneously in a time series:
  - A trend in the mean
  - A seasonal component that shows a regular pattern repeating on a set cycle (yearly or daily temperatures for example)
  - A cyclic component that is separate from seasonality (19 year tidal cycle)
  - Excursions due to an outside influence (may be temporary or permanent)
  - A random error component

Components of a Time Series



Module 8.2

Components of a Time Series



- This series has a linear trend
- and a 12 month seasonal component
- and some sort up upset at 37 months
- and random error

Module 8.2

Components of a Time Series

- Any given time series may have a combination of these sort of characteristics.
- The main thing about time series analysis is that it is needed when the data are autocorrelated - that's when the standard techniques fail.

Module 8.2

Detection of Change Points

- Detecting a change point means finding an upset or shift in the mean.
- If it happened at a known time (a known release occurred) then it involves testing for a change in the mean before and after the event.
- Note that you must use techniques that do not assume independence if significant autocorrelation exists.

Module 8.2

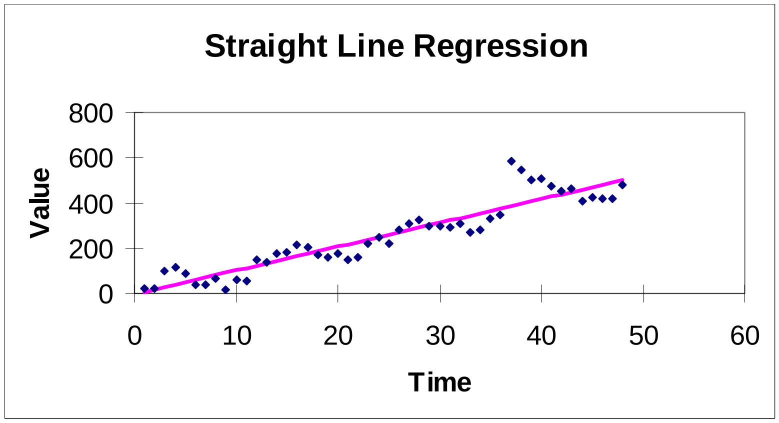
Detection of Change Points

- Nonparametric techniques will work in this case.
- If the change occurred at an unknown time, it is a very difficult problem. Consult an expert.

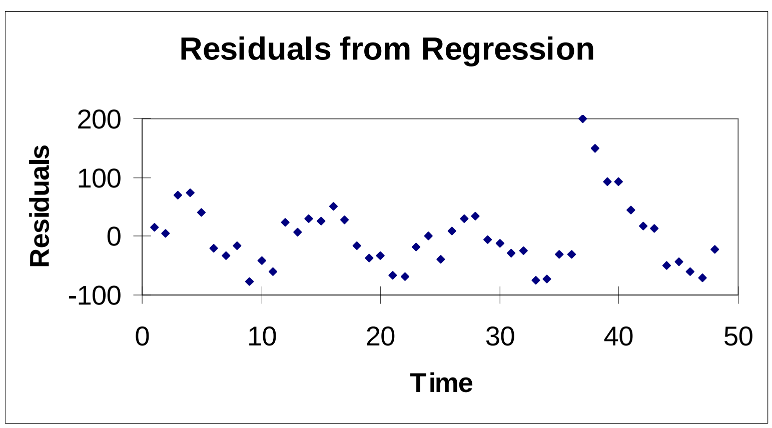
Detection of a Trend

- Another common problem is detecting and modeling a trend.
- With independent data points, use simple least squares regression analysis.
- After the model is fit, test for autocorrelation in the residuals using the Durbin-Watson statistic
- If the residuals are autocorrelated, you must use a different method of fitting than least squares or use time series methods

Detection of a Trend



Detection of a Trend



Detection of Autocorrelation

- Durbin-Watson Test

$$V = \frac{\sum_{i=2}^{n}(e_i - e_{i-1})^2}{\sum_{i=1}^{n} e_i^{2}}$$

- where e_i to e_n are the residuals from the regression analysis
- V will be 2 if there is no autocorrelation, significantly more than 2 if observations are negatively autocorrelated, and less than 2 if positively autocorrelated

Detection of Autocorrelation

- Durbin-Watson Test
  - Table A2.5 (Manly p 278) gives the critical values for the Durbin-Watson Test
  - For example, for a regression of Y on X with 20 data points, run the regression, obtain the residuals, and calculate V.
  - If positive autocorrelation exists (most common in environmental data), V would be less than 2. The data are definitely significantly positively autocorrelated if V<1.08.

Detection of Autocorrelation - Durbin-Watson Test Example

Time	Predicted Y	Residuals	Square of Residuals	Differences	Square of Differences
1	8.02	15.44	238.52		
2	18.51	3.87	14.99	-11.57	133.93
3	29.01	69.90	4885.66	66.03	4359.46
4	39.50	74.11	5492.77	4.22	17.77
5	49.99	40.23	1618.31	-33.88	1148.19
6	60.48	-20.48	419.41	-60.71	3685.44
7	70.97	-33.38	1114.54	-12.91	166.54
8	81.46	-16.43	269.80	16.96	287.61
9	91.95	-77.01	5930.29	-60.58	3670.28
10	102.44	-42.07	1770.22	34.93	1220.41
11	112.94	-59.35	3521.98	-17.27	298.33
12	123.43	23.27	541.37	82.61	6825.03
13	133.92	6.43	41.41	-16.83	283.34
14	144.41	30.59	935.79	24.16	583.50
and so on	and so on	and so on	and so on	and so on	and so on
			155675.64		**108745.64**
				V =	**0.70**

Module 8.2

13

Detection of Autocorrelation - Durbin-Watson Test Example

- Null Hypothesis: Data are independent
- Alt. Hypothesis: Data are autocorrelated
- V=0.70
- n=48
- p=1 (regression on one variable)
- d_1 = 1.42 (for n=50)
- d_2 = 1.50 (for n=50)
- The test is significant since V is less than 1.42. Reject the null hypothesis.

Module 8.2

14

Detection of Autocorrelation - Durbin-Watson Test Example

- Conclusion: The data are significantly positively autocorrelated. The use of regression analysis is inappropriate since regression analysis assumes independent data
- (Note on Durbin-Watson test: If the test had been close, use linear interpolation to calculate the values for d_1 and d_2 in between the values given in the table)

Module 8.2 15

Detection of a Trend

- It's becoming more and more common to use nonparametric tests of trend for environmental data:
  - Mann-Kendall Test
  - Seasonal Kendall Test
  - Moving averages
- These tests do not assume normally distributed errors; However, the first two still assume independent data points
- Null hypothesis is always no trend exists

Module 8.2 16

Detection of a Trend

- Mann-Kendall Test
  - Use for data that do not have a seasonal component

$$S = \sum_{i=2}^{n} \sum_{j=1}^{i-1} sign(x_i - x_j)$$

  - where $sign(x_i - x_j)$ is
    - -1 for $x_i - x_j < 0$
    - 0 for $x_i - x_j = 0$
    - 1 for $x_i - x_j > 0$

Detection of a Trend

- Mann-Kendall Test
  - If no trend exists, then S is expected to be zero and has variance

$$Var(S) = n(n-1)(2n+5)/18$$

  - If n>10, compare the test statistic

$$Z = S/SE(S)$$

  - against the critical values from the Standard Normal Table

Detection of a Trend

- Mann-Kendall Test Example

5										
6	1									
8	1	1								
7	-1	1	1							
6	-1	-1	0	1						
7	1	0	-1	1	1					
8	1	1	1	0	1	1				
9	1	1	1	1	1	1	1			
9	0	1	1	1	1	1	1	1		
10	1	1	1	1	1	1	1	1	1	
Sums =	**4**	**5**	**4**	**5**	**5**	**4**	**3**	**2**	**1**	**33**

Detection of a Trend

- Mann-Kendall Test Example
  - S = 33
  - Var(S) = n(n-1)(2n+5)/18 = 10*9*25/18=125
  - SE(S) = 11.2
  - Z = S/SE(S) = 33/11.2 = 2.95
  - Prob (Z<-2.95 and Z>2.95)=2*0.002 = 0.004
  - Conclusion: Reject the null, there is a trend

Detection of a Trend

- Seasonal Kendall Test
  - Use for data that do have a seasonal component
  - Calculate S separately for each season
    - A season could be a month or could be groups of months like Spring, Summer, Fall, Winter
    - Sum all of the S's to get S_T
  - S_T has expected value zero and variance equal to the sum of the variances of the S's

Module 8.2

21

Detection of a Trend

- compare the test statistic

$$Z = S/SE(S_T)$$

against the critical values from the Standard Normal Table

Module 8.2

22

Time Series Models

- Throughout this module, I have referred to time series analysis but haven't described what it is exactly.
- Much of that is because time series analysis is complex and it's details are beyond the scope of this course
- However, I will give a brief overview of time series analysis

Module 8.2 20

Time Series Models

- A time series model involves explicitly modeling the residuals as autocorrelated
- For example:
  - $\varepsilon_t = \alpha\varepsilon_{t-1} + u_t$ or
  - $\varepsilon_t = \alpha_1\varepsilon_{t-1} + \alpha_2\varepsilon_{t-2} + u_t$
  - and so on

  where u_t are normally distributed random errors

Module 8.2 21

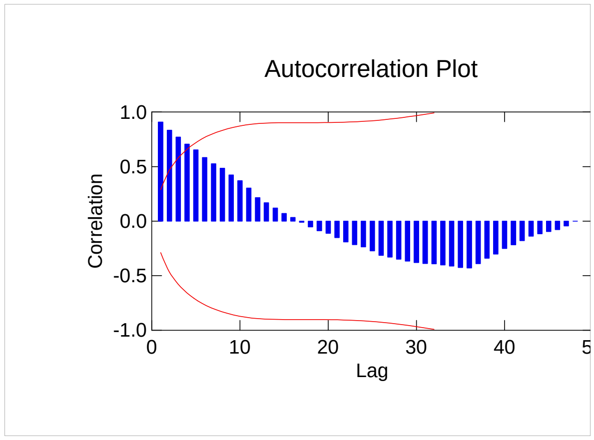
Time Series Models

- There are several types of models for the errors:
  - Autoregressive (AR)
  - Moving average (MA)
  - Autoregressive moving average (ARMA)
  - Integrated autoregressive moving average (ARIMA)

Time Series Models

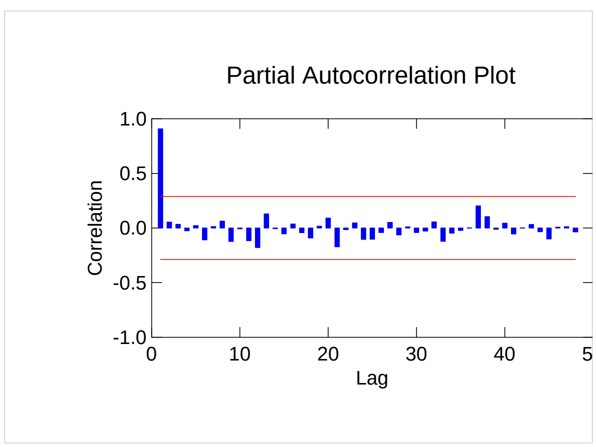
- Special software is needed to help determine which of these models is best for a series
- Two correlograms are used: the autocorrelation function (ACF) and the partial autocorrelation function (PACF)
- Together they can help the modeler determine if the best error model is AR, MA or a mixed model

Time Series Models



Module 8.2

Time Series Models



Module 8.2

Time Series Models

- Time Series Steps:
  - Differencing to remove trend
  - Seasonal differencing to remove seasonality
  - Re-do ACF and PACF plots
  - Determine order of AR and MA portions of series
  - Use to forecast

Module 8.2

Time Series Analysis

- The main point of this module is to be aware that autocorrelation can exist when data points are close together in time and that the presence of autocorrelation can make the use of standard techniques such as regression analysis inappropriate.
- Alternative techniques include nonparametric analysis and time series analysis

Module 8.2