# TIME SERIES

**by Michele MICCIO**

**Dept. of Industrial Engineering (Università di Salerno)**

**Prodal Scarl (Fisciano)**

rev. 3 of November 27, 2013

## Time Series: Definitions

A **time series** associates each data element with a specific point in time.

*Frank Smietana* in Mathcad documentation

Values taken by a variable over time (such as daily sales revenue, weekly orders, monthly overheads, yearly income) and tabulated or plotted as chronologically ordered numbers or data points.

http://www.businessdictionary.com/

A **time series** is a sequence of data points, measured typically at successive times spaced at uniform time intervals.

$Y_1$, $Y_2$,…, a series of values measured or observed for a variable Y at successive times $t_1$, $t_2$,… which may be reported as a sequence of integers $k = 0, 1, 2, . . .$

2

ⓘ **time series are NOT just data …**

# Time Series: Examples

- the daily closing value of the Dow Jones index
- the annual flow volume of the Nile River at Aswan
- the daily temperature readings
- the sampled pressure measurements in a gas-phase chemical reactor
- the daily production of pasteurized milk plastic bottles in a factory

. . .



# Time Series Representation

| TEMPO | VALORI |
|---|---|
| 2000 | 100 |
| 2001 | 120 |
| 2002 | 98 |
| 2003 | 100 |
| 2004 | 110 |
| 2005 | 115 |
| 2006 | 120 |

Table representation


Graphical representation

(Ex. of a line diagram)

ⓘ

- **A time series is usually stored and made available in a computer file**
- **The time points are typically, but not necessarily, equidistant. For example, the daily factory output values do not contain data for holidays and weekends**

## Time Series classification

- **deterministic**
- **stocastic or probabilistic**
- **mixed (deterministic and stocastic together)**

e.g., of **additive type**:

$$y_k = f_k + a_k$$

where:

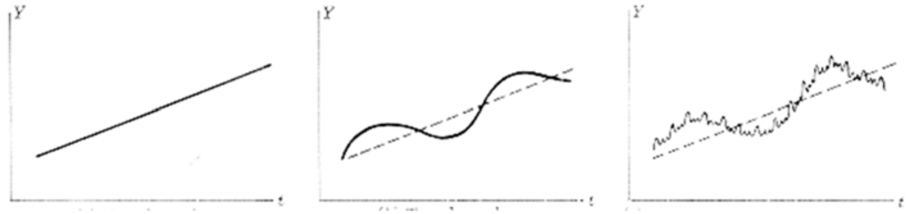
$y_k$ = time series value at time $k$

$f_k$ = deterministic component at time $k$

$a_k$ = probabilistic component at time $k$

5

## Time Series patterns
(mainly in economics)



Long term Trend | Long term Trend + cyclic pattern | Long term Trend + cyclic pattern + seasonal pattern + random fluctuations

Time series consist of four components:

Trend variations that move up or down in a reasonably predictable pattern,

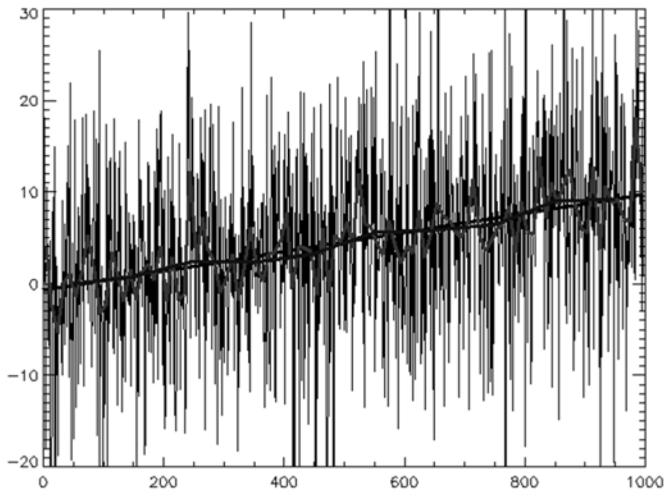
Cyclical variations that correspond with business or economic 'boom-bust' cycles or follow their own peculiar cycles, usually multi-annual

Seasonal variations that repeat over a specific period such as a day, week, month, season, etc.,

Random variations that do not fall under any of the above three classifications.

6

# Time Series patterns




Ex. of a time series with:
- random data (black oscillations)
- trend (black line),
- best-fit line (blue)
- and smoothing (red curve)



# **Time series analysis**

Time series analysis comprises methods for

1. calculation/extraction of meaningful statistics (mean, variance, symmetry, kurtosis, autocorrelation, etc.)
2. *Trend* determination, with possible suppression of the oscillating pattern (filtering)
3. Estimation of cyclic patterns and their period
4. Estimation of seasonal variations and their period
5. *Detrending* the time series
6. Model identification for subsequent simulation and prediction
7. Estimation of missing or lost data in the series
8. Analysis/search of deterministic chaos
   - e.g. determination of "*Lyapunov* exponents "

Examples in engineering applications:

- signal filtering
- design of predictive controllers

## Time series analysis: 2. *trend* determination

1. **least squares**
2. **empirical modeling**
3. **. . .**



# Time series analysis: *Example No. 1*

Time Series Regression of Airline Passenger Data

Demo and data available in Matlab

Air Passenger Data

an array of monthly counts of airline passengers, measured in thousands, for the period January 1949 through December 1960:

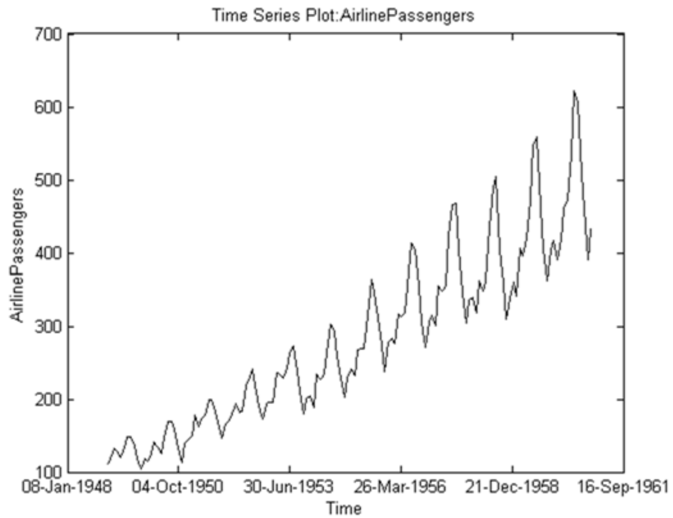
```
%   1949 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960
y = [112  115  145  171  196  204  242  284  315  340  360  417    % Jan
     118  126  150  180  196  188  233  277  301  318  342  391    % Feb
     132  141  178  193  236  235  267  317  356  362  406  419    % Mar
     129  135  163  181  235  227  269  313  348  348  396  461    % Apr
     121  125  172  183  229  234  270  318  355  363  420  472    % May
     135  149  178  218  243  264  315  374  422  435  472  535    % Jun
     148  170  199  230  264  302  364  413  465  491  548  622    % Jul
     148  170  199  242  272  293  347  405  467  505  559  606    % Aug
     136  158  184  209  237  259  312  355  404  404  463  508    % Sep
     119  133  162  191  211  229  274  306  347  359  407  461    % Oct
     104  114  146  172  180  203  237  271  305  310  362  390    % Nov
     118  140  166  194  201  229  278  306  336  337  405  432 ]; % Dec
% Source:
% Hyndman, R.J., Time Series Data Library,
% http://www-personal.buseco.monash.edu.au/~hyndman/TSDL/.
% Copied in October, 2005.
```

ⓘ **Tabular representation of a time series**

# Time series analysis:
# *Example No. 1*

Time Series Regression of Airline Passenger Data

Demo and data available in Matlab




ⓘ **Graphical representation of a time series**

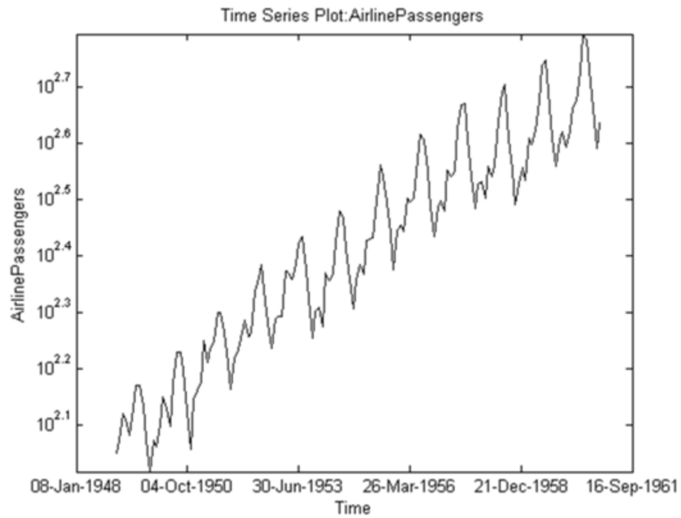


# Time series analysis:
# *Example No. 1*

Time Series Regression of Airline Passenger Data

Demo and data available in Matlab

Switch from linear to log scale on y-axis



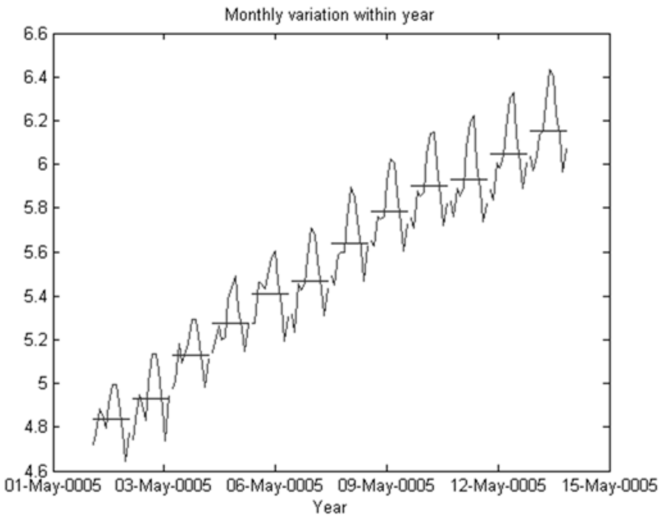

ⓘ **Graphical representation of a time series**

# Time series analysis: *Example No. 1*

Time Series Regression of Airline Passenger Data

Demo and data available in Matlab

Monthly within years




ⓘ **the month-to-month variation within years appears constant.**

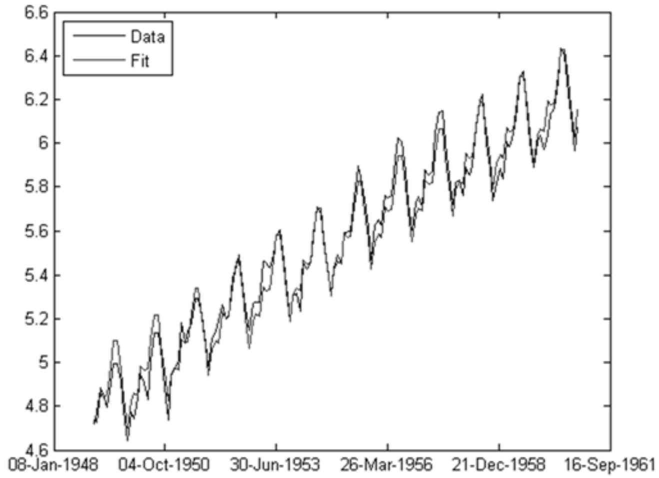
13

# Time series analysis: *Example No. 1*

Time Series Regression of Airline Passenger Data

Demo and data available in Matlab

Fitting model to give a linear trend plus a cyclical component




14

# Time series analysis:
# 2. *Filtering*

Objectives:

Eliminate fluctuations:

- eliminate cyclic, seasonal and irregular patterns
- allow *trend* determination

Investigate fluctuations

Methods:

- **moving average**
- **Fourier transform**
- . . .



## Time series filtering
### Moving Average of order M

Original Time Series:

$Y_1 , Y_2 , Y_3 \dots , Y_i \dots Y_N$

New Time Series generated by a simple moving average (using a statistical tool, e.g., mean, median, etc., or a user-defined weighting scheme):

$Y_{ma,1} , Y_{ma,2} \dots , Y_{ma,i} \dots Y_{ma,N}$

As an example, with arithmetic mean:

$$\mathbf{Y_{ma,i} = (Y_{i-M/2} + \dots + Y_i + \dots + Y_{i+M/2})/(M+1)}$$

with the span $M \in \mathcal{N}$ and even

ⓘ The new Time Series turns out smoothed.

# Filtering Algorithm with Moving Average

1) let's have a time series with N $\in \mathcal{N}$ points and let be M $\in \mathcal{N}$ (even) << N

2) calculate $Y_{new,i}$ from $Y_i$ with arithmetic mean from i=M/2+1 to i=N-M/2-1

ⓘ the new $Y_{new,i}$ series contains the trend and low frequency **oscillations only**.

3) construct a third series $\mathbf{Y_i^*}$ by subtraction:

$$\mathbf{Y_i^* = Y_i - Y_{new,i}}$$

ⓘ the new $Y_i^*$ series contains the high frequency **oscillations only**.

ⓘ

The first as well as last M/2 points in the original series have to be discarded

The span M is a smoothing parameter:

- Depending on the span chosen, either high or low periodicity is revealed by the smoothed series
- M=0 gives the original series back
- M=N (in principle, M=largest even No.<N) gives the overall arithmetic mean



# Example No.2

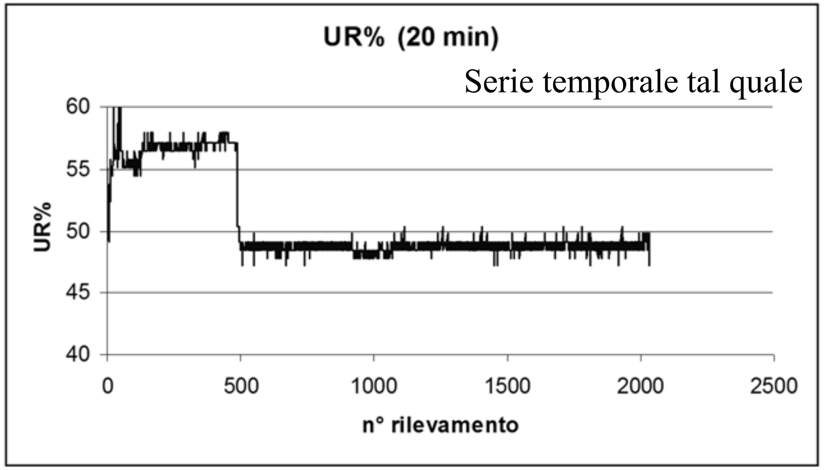
Time series filtering with Moving Average of order M
During salami "Nebrodi" ripening

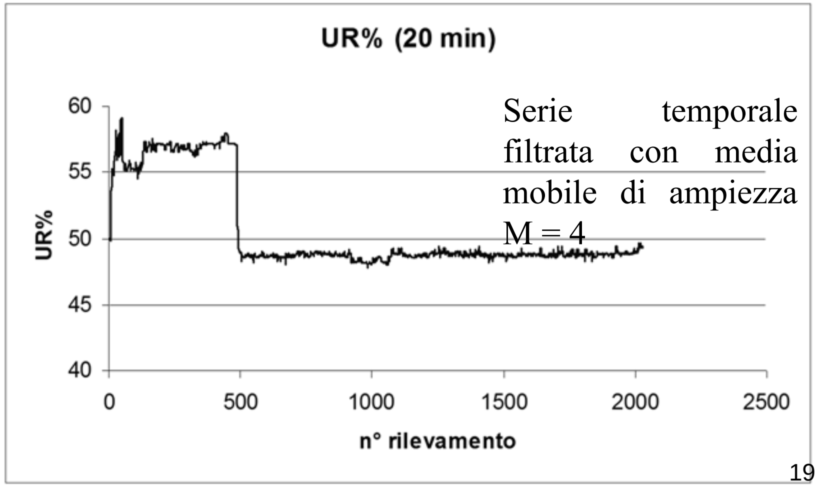
Stazione Sperimentale per l'Industria delle Conserve Alimentari (Parma/Angri)

🕮 **Tesi Rendina 2005-06**

# Time series of RH data

sampling every 20 min (n° rilevamento)

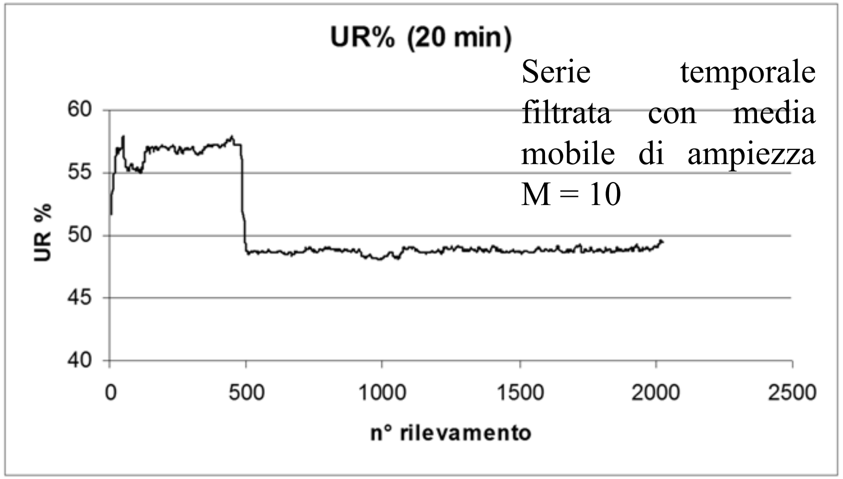


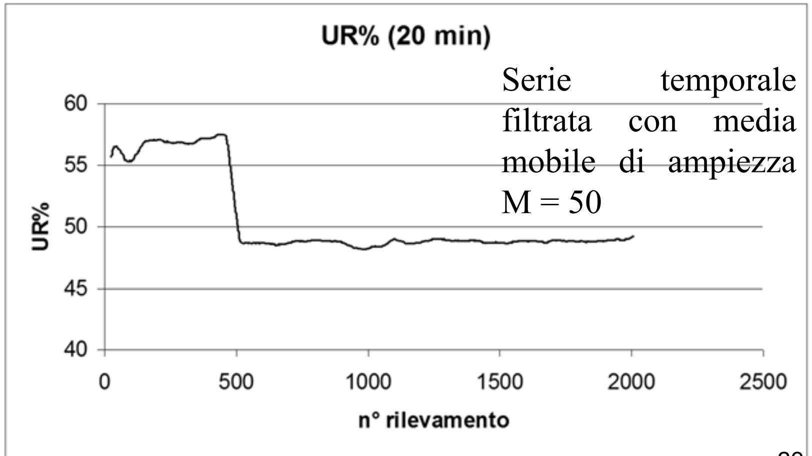






# Time series of RH data

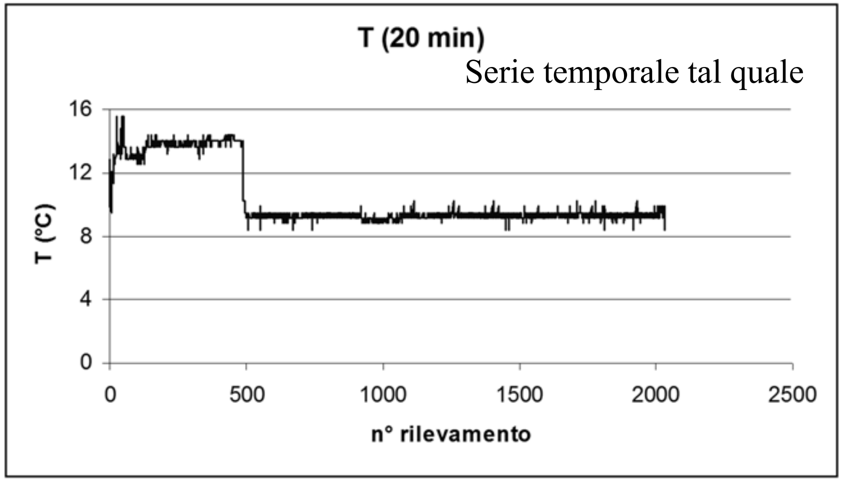
sampling time 20 min

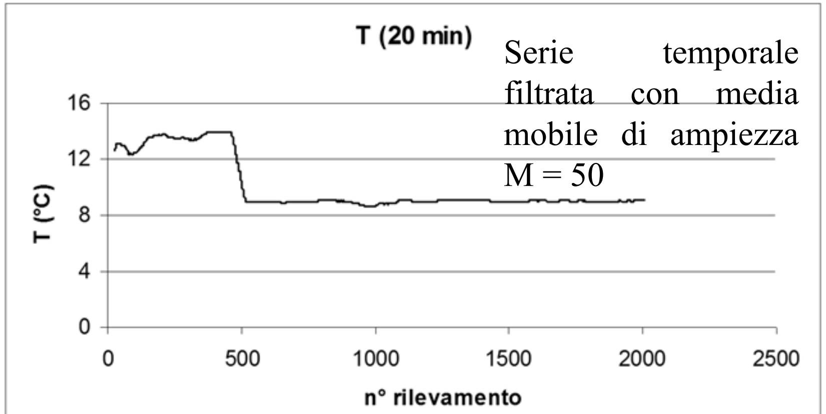


# Time series of T data

sampling time 20 min






# Time series analysis: 3. *Detrending*

Definition:

Detrending is the statistical or mathematical operation of removing trend from the series.

Objective:

Detrending is often applied to remove a feature thought to distort or obscure the relationships of interest.

For example, in climatology a temperature trend due to urban warming might obscure a relationship between cloudiness and air temperature.

Detrending is also sometimes used as a preprocessing step to prepare time series for analysis by methods that assume stationarity.



http://www.ltrr.arizona.edu/~dmeko/geos585a.html

# Time series analysis: 3. *Detrending*

Methods:

Simple linear trend in mean can be removed by subtracting a least-squares fit straight line from the series.

More complicated trends might require different procedures. Four alternative approaches to detrending are:

- **first differencing**

A time series that is non-stationary in mean (e.g., trend in mean) can be made stationary by taking the first difference. The first-difference is the time series at time t minus the series datum at time t – 1 :

$w_i = x_i - x_{i-1}$

where $w_i$ is the first-differenced series.

- **curve-fitting**
- **digital filtering**
- **piecewise polynomials** (e.g., cubic smoothing spline)

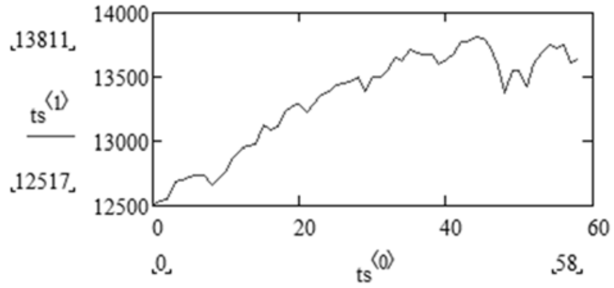


http://www.ltrr.arizona.edu/~dmeko/geos585a.html

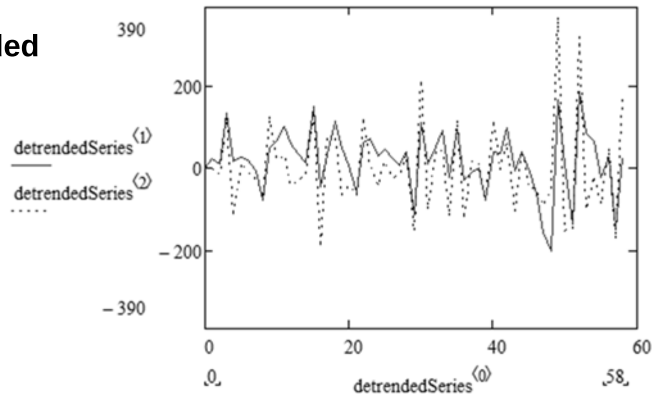
# Time series analysis: 3. *Detrending*

**EXAMPLE**

**Original Time Series**




**Detrended Time Series**




*Frank Smietana* in Mathcad documentation

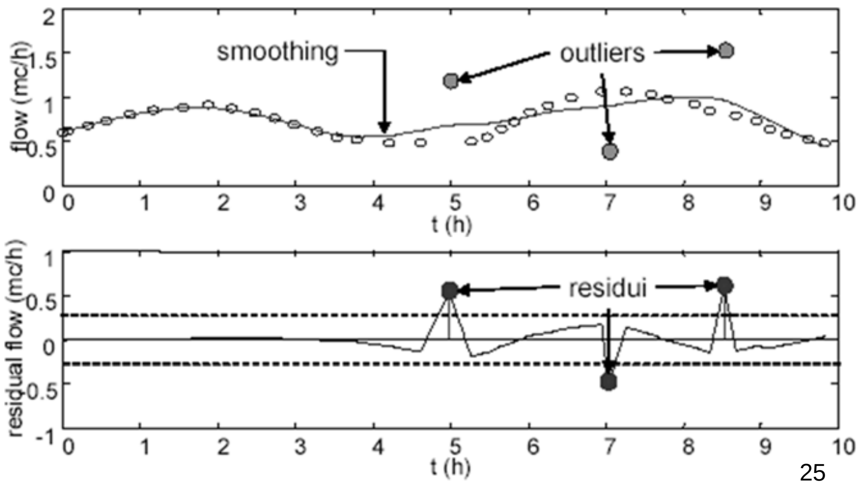
# *Outliers*

Definition:

**outliers** are data points that are part of the series, but "apparently" are beyond an admissibile *range*

Methods for searching *outliers*:

- Filtering or *smoothing* data with a simple *fitting* model
- Calculating residuals as differences between the original value and the filtered (or the *fitting* model) value
- o *Outliers* are just the original data whose residuals go beyond a predefined threshold






# Time series analysis:
# 4. *Forecasting*

## Definition:

Time series forecasting is the use of a model to predict future values based on previously observed values.

a. short term *trend*

b. long term *trend*

## Methods:

- Extrapolation from a *fitting* model
- Development of a *black box* model (e.g., ARMA model)
- . . .