Time Series Workshop

minhtran

Apr 2016

Agenda

- Introduction to Time Series
- Time Series Representation
- Time Series Characteristics
- Time Series Modeling & Demo
- Discussion
- Conclusion

Agenda

- Introduction to Time Series

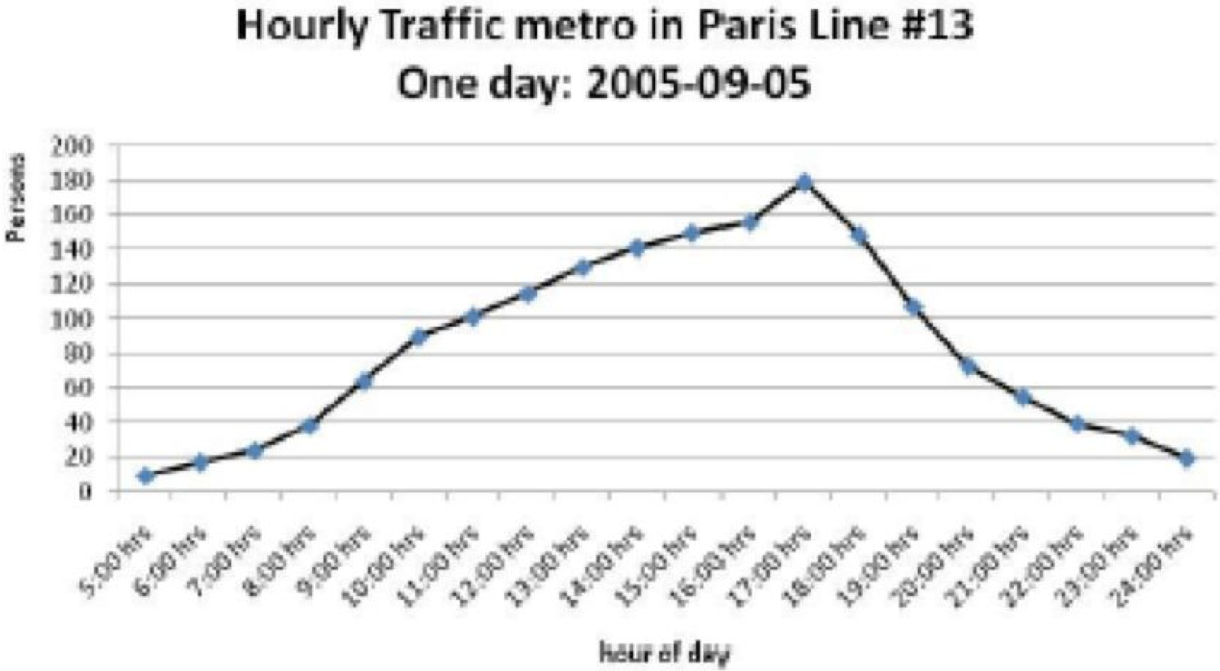
Introduction to Time Series

- Why time series is important?
  - Predict/Forecast
    - Because most of business and projects require some planning, which most of the time is performed with an uncertainty knowledge of future conditions
    - Because it is mandatory to measure the possible risks around future events
    - Because most of the time it is required to calculate some metric indices, which may be related to economy, politics, technology etc.
  - Evaluation
    - Because designs of new systems need to be evaluated carefully before the systems are manufactured
    - Evaluation needs as much data as possible to be realistic
    - But data is precious and a data set is only one data point
    - More data set can be generated by models

Time Series

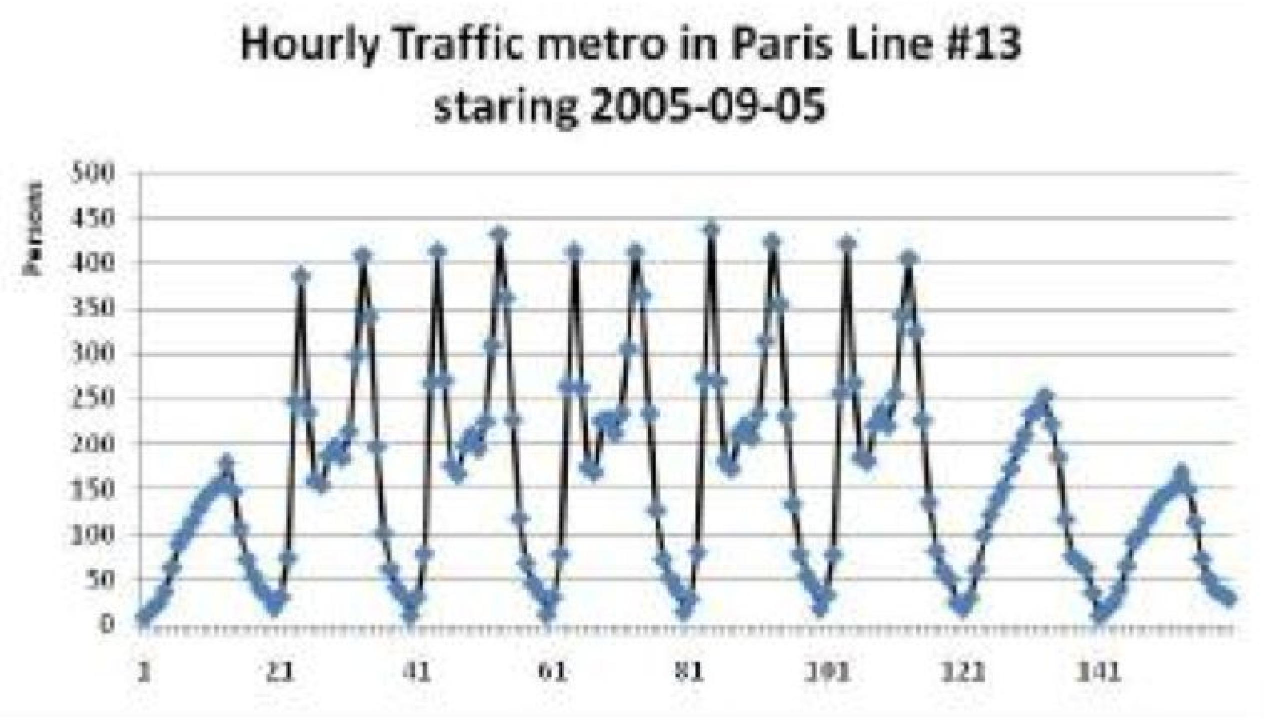
- A time series is a signal that is measured in regular time steps
- A time series is also referred to as a stochastic process such as a point process, an arrival process, an interarrival process, a count process, a rate process, etc.
- The time step of a series may be of any length like seconds, minutes, hours, days, weeks, months, years, etc. This will bring on very different “looks” of the time series
- The time step can also be called time resolution, time interval, time scale, etc.

Example: A Time Series Measured each Hour



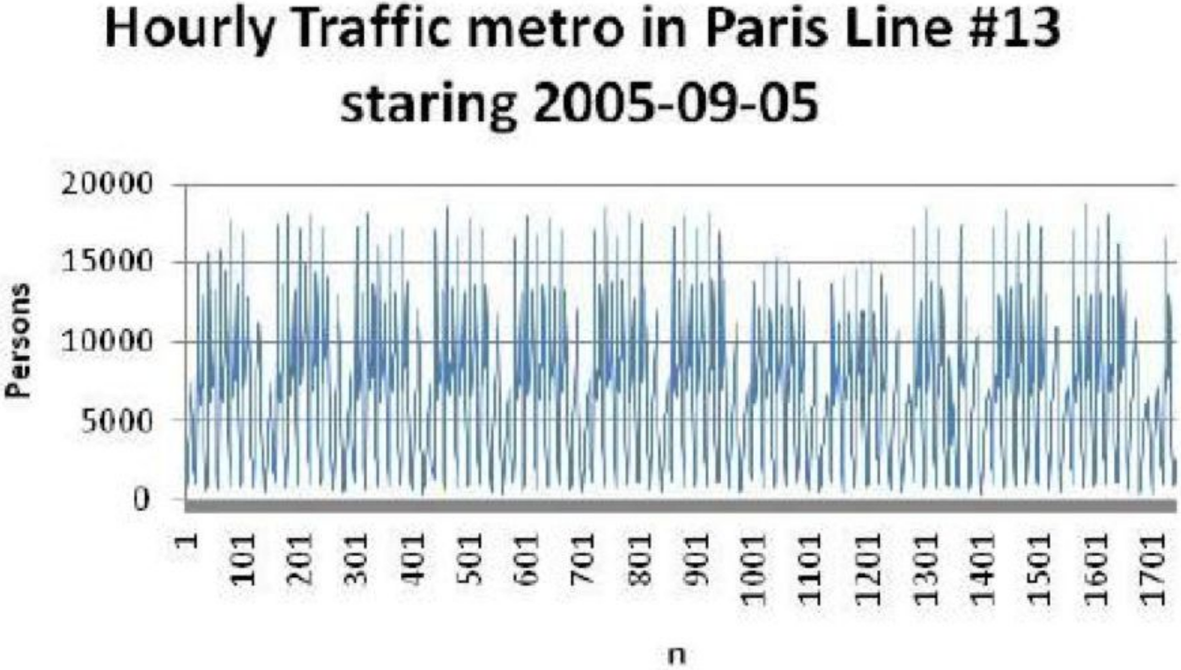
INAOE 2014. P. GOMEZ-GIL et al.

Few Days of the Same Time Series



INAOE 2014. P. GOMEZ-GIL et al.

Few Months of the Same Time Series



INAOE 2014. P. GOMEZ-GIL et al.

Agenda

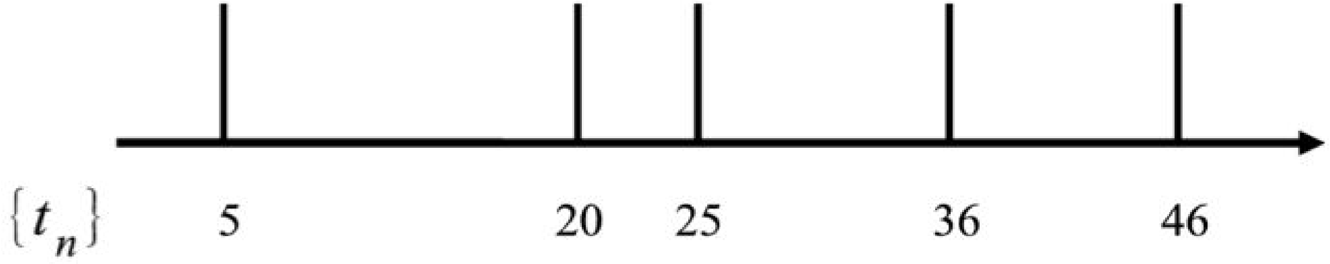
- Introduction to Time Series
- Time Series Representation

A Sample of Time Series Data

_user	_item	_rating	_time
17860	1004	3	2004-08-15
1516313	1004	3	2004-08-18
387418	1004	1	2004-08-21
1559083	1004	3	2004-08-21
305344	1004	1	2004-08-21
642384	1004	5	2004-08-23
572481	1004	1	2004-08-31
888162	1004	2	2004-08-31
1416673	1004	2	2004-09-01
2116582	1004	4	2004-09-02
2253294	1004	3	2004-09-02
2214785	1004	3	2004-09-02
1349711	1004	4	2004-09-03
2213550	1004	3	2004-09-03
1487271	1004	2	2004-09-03
194528	1004	3	2004-09-03
706897	1004	3	2004-09-03
323071	1004	4	2004-09-03

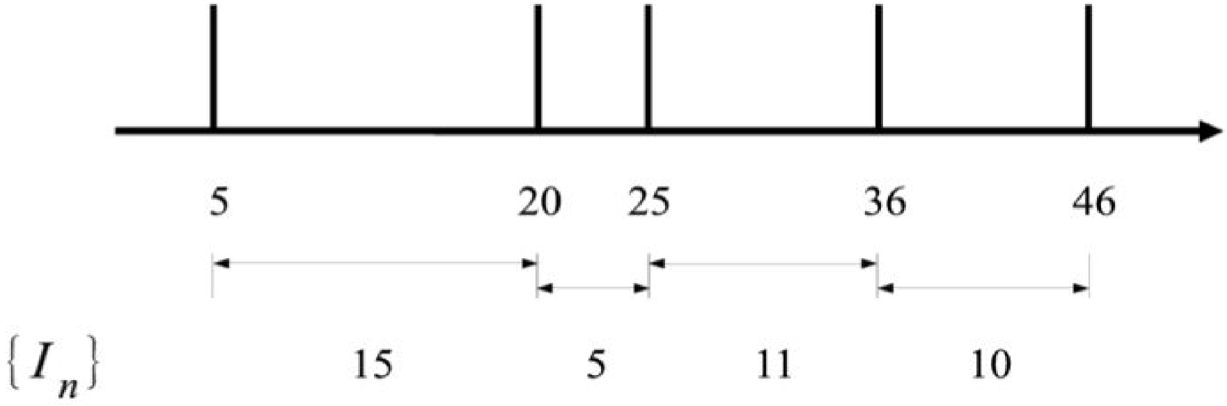
Arrival (time) process

Representation of Arrivals - Arrival Process



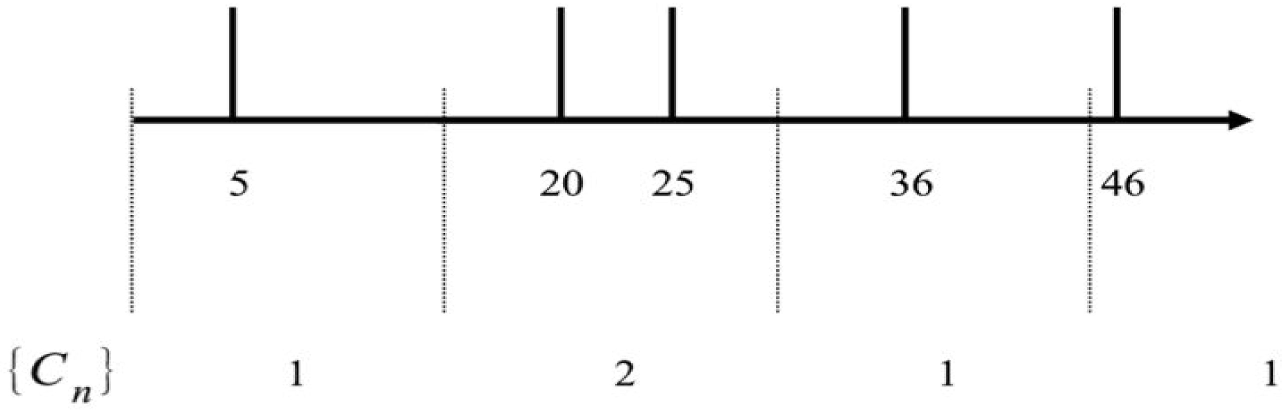
- A natural representation of arrival events
- But rarely used in practice

Representation of Arrivals - Interarrival Process



- Keep the whole information about arrivals and so can be used to recreate the original point process accurately
- The direct correspondence between its index number and the absolute time is lost

Representation of Arrivals - Count Process



- Need to define a time interval parameter T
- The direct correspondence between its index number and the absolute time is kept
- Information about the times between arrival events within an interval is lost

Representation of Arrivals - Rate Process

- Same advantages and disadvantages as a count process
- Most of magics happen with this type of representation

Agenda

- Introduction to Time Series
- Time Series Representation
- Time Series Characteristics

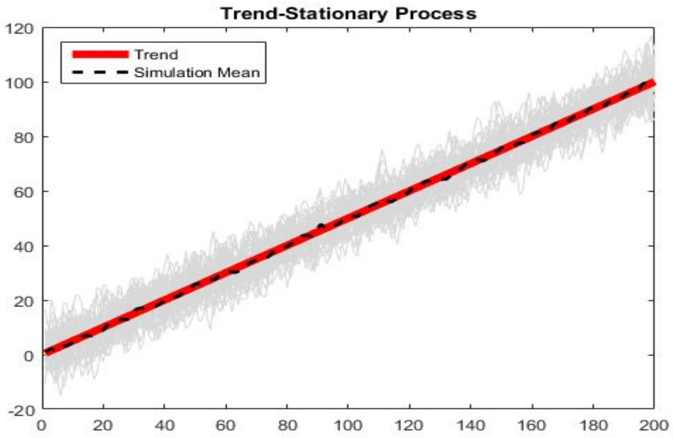
Methodology to Analyse a Time Series

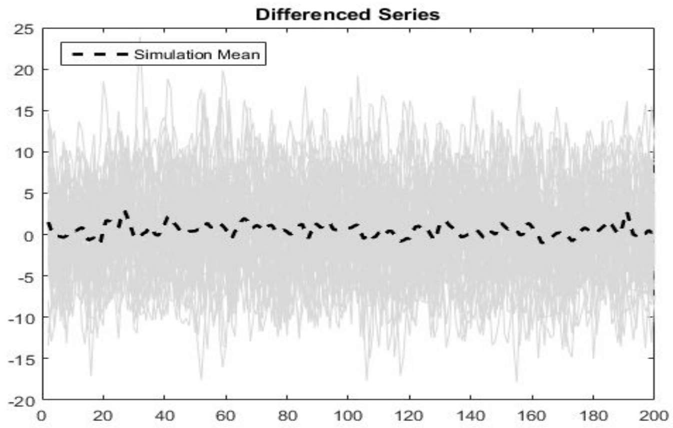
- Start with a simple statistical analysis
  - Five box statistics, std, variance
- Look at the distribution of the time series
  - Histogram, CDF, CCDF
- Analyse the characteristics of the time series
  - Stationary
  - Periodicity
  - Trend
  - Cross-correlation
  - Auto-correlation
  - Long range dependence
  - Burstiness
  - Temporal locality

Methodology to Analyse a Time Series

- Some approaches
  - Hypothesis testing
  - Periodicity transform
  - Wavelet transform
  - Fourier transform
  - Measure the Hurst parameter
  - Entropy
  - etc.

Stationary

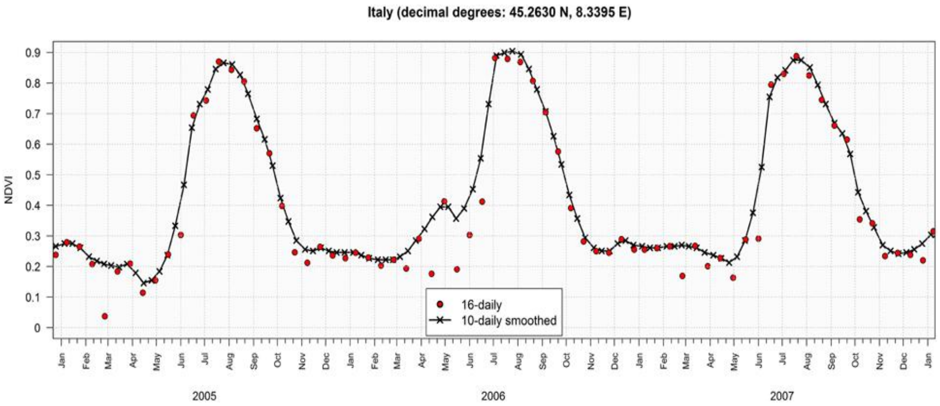


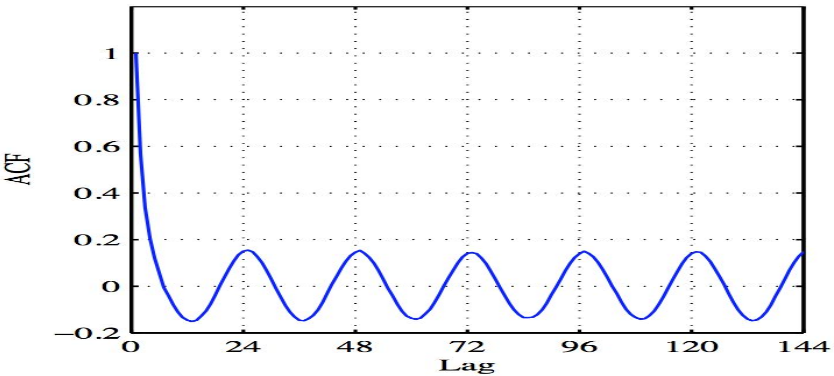


Source: www.mathworks.com

- Can be detected by using hypothesis testing such as the kpss test

Periodicity (Seasonal)

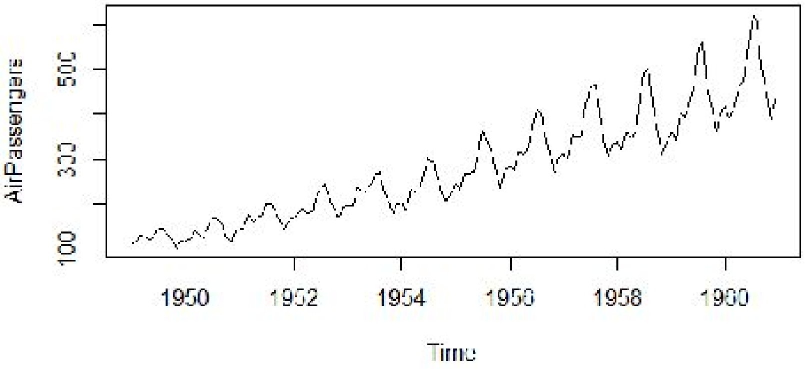


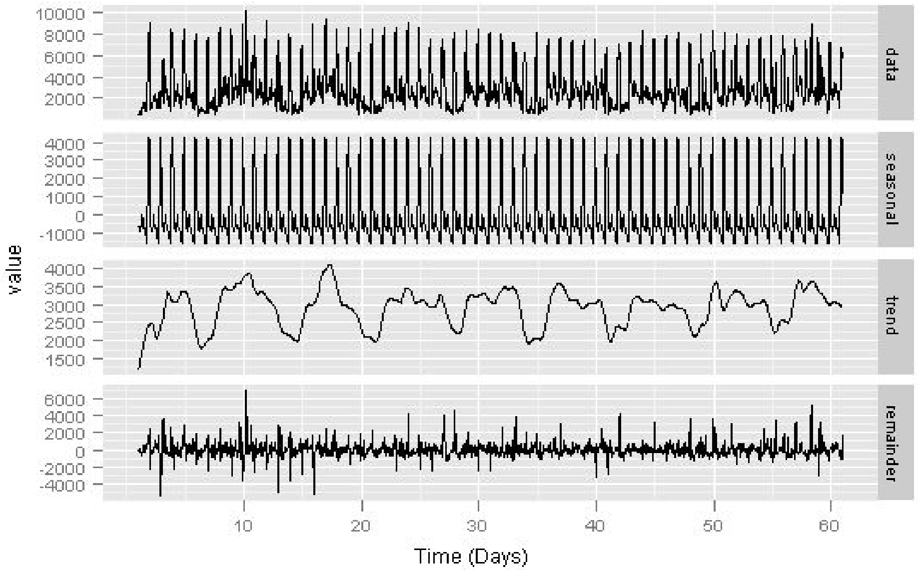


Source: http://journal.frontiersin.org

- Can be detected by using hypothesis testing such as the fisher.g.test test
- Can be easily observed by drawing the ACF function
- Can be decomposed by the periodicity transform theory

Trend



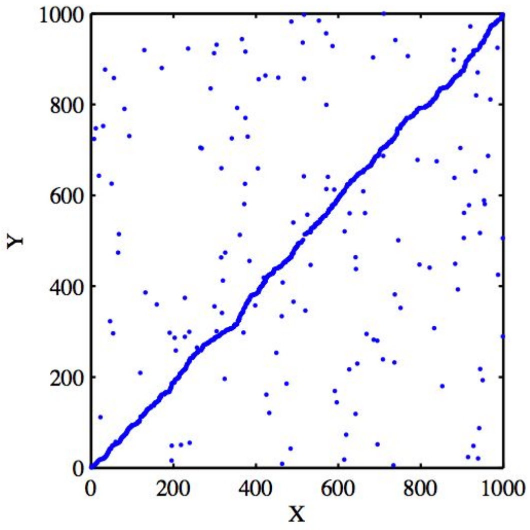


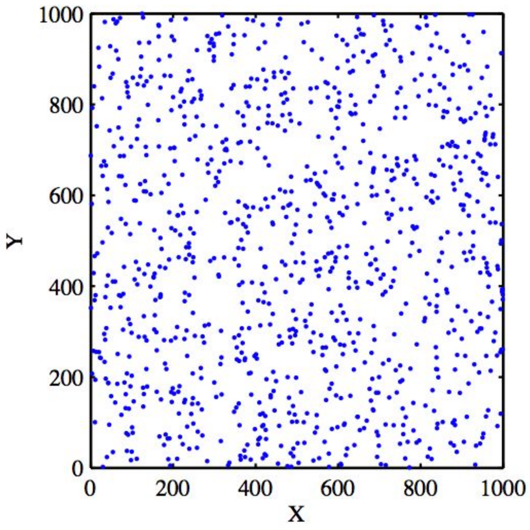
Source: www.quora.com, https://learnr.wordpress.com

- Can be detected by using hypothesis testing such as the Mann-Kendall test

Cross-correlation

$$\rho_{X,Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y}$$





- Can be quantitatively measured by Pearson's, Spearman's, Kendall's or distributional correlation

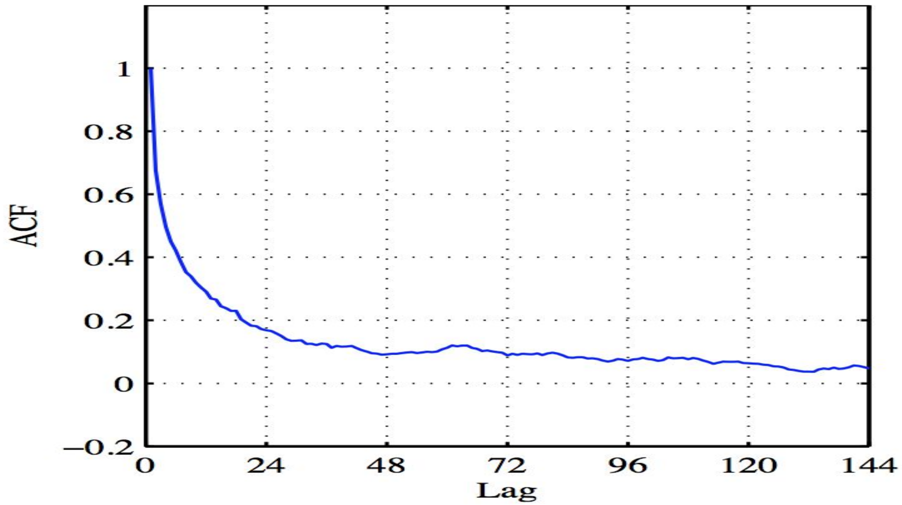
Auto-correlation

The autocorrelation function (ACF) of a stochastic point process $X = \{X_n\}$ describes the correlations between values of X at different points in time [48]. If X is second-order stationary, its ACF is defined as

$$R(k) = \frac{E[(X_i - \mu)(X_{i+k} - \mu)]}{\sigma^2}, \tag{2.1}$$

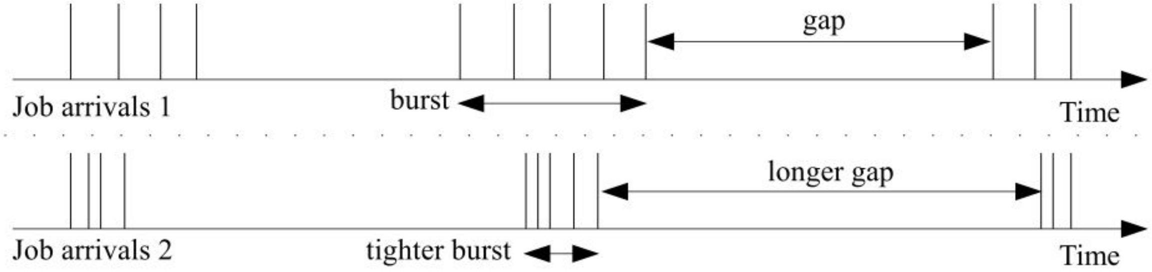
where $E[\cdot]$ is the expected value, k is the time shift being considered (usually referred to as the lag), and μ and σ^2 are the mean and the variance, respectively, of X.

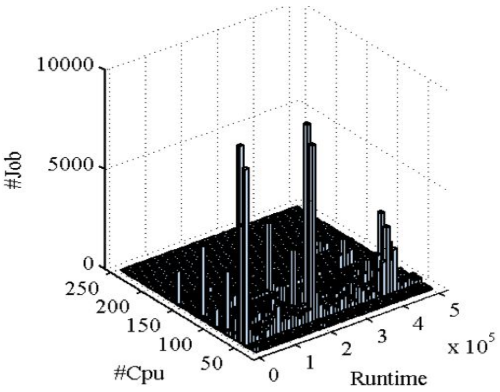
Long Range Dependence

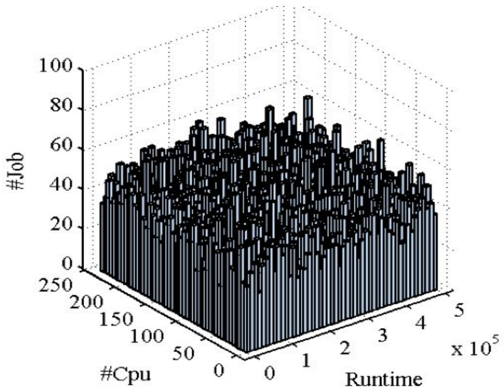


- Can be quantitatively measured by estimating the Hurst parameter
- Approaches include:
    - *Absolute Value*
    - *Aggregate Variance*
    - *R/S*
    - *Variance of Residuals*
    - *Periodogram*
    - *Whittle*
    - *Abry-Veitch*

Burstiness







- Temporal burstiness is measured by the coefficient of variation of interarrivals

- Spatial burstiness is measured by the normalized entropy

Temporal locality

- The phenomenon of temporal locality is understood as a persistent similarity between consecutive values within a time series
- Considered a time series R = {10, 12, 15, 9, 3000, 2800, 400, 360, 420}, assume we have an efficient approach to classify R in such a way that similar runtimes are grouped to the same cluster

  → a series of cluster labels L = {A, A, A, A, C, C, B, B, B}

  → R has temporal locality

Agenda

- Introduction to Time Series
- Time Series Representation
- Time Series Characteristics
- Time Series Modeling & Demo

Time Series Modeling

- Modeling problems
  - Synthetic modeling (the so-called)
    - Create/build models based on train data and then use the models to generate synthetic data
    - Example: language modeling, workload modeling in systems or networks
  - Predictive modeling
    - Create/build models based on train data and then use the models to predict future values
    - Example: regression/classification problems

⟶ Deep learning can be applied to both

Time Series Modeling

- Modeling types
  - Single time series
    - The estimation of future values in a time series is commonly done using past values of the same time series
    - Synthetic modeling often targets at capturing the characteristics of the input time series
  - Multiple time series
    - “Normal” machine learning techniques + some special “tricks” to deal with the time factor
    - Synthetic modeling often targets at capturing the relationship among several time series such as the cross-correlation and spatial burstiness

Time Series Modeling

- 2 modeling problems x 2 modeling types = 4 combinations for classifying a learning problem

Synthetic Modeling on Single Time Series

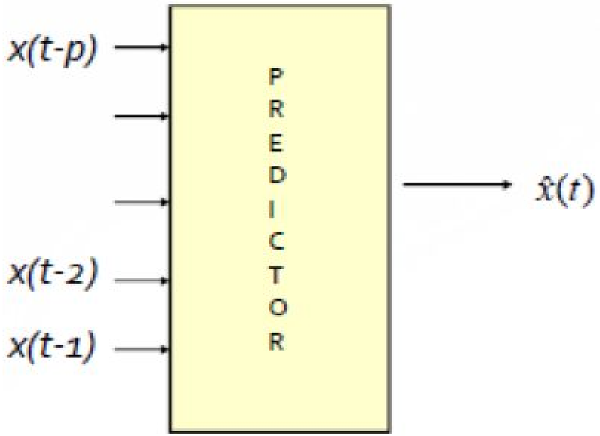
- Model probability distribution
  - Use MLE to estimate parameters
  - Use non-parametric methods
  - Validated by hypothesis testing like KS-test, KS2-test, AD-test, etc.
- Model time series characteristics
  - The multifractal wavelet model to model the dependency
  - The periodicity transform theory to model the periodicity
  - The model-based clustering to model temporal locality
- Model hidden states
  - Hidden Markov Model
- Usually a favourite model can capture well several features at the same time together with the probability distribution
- Demo

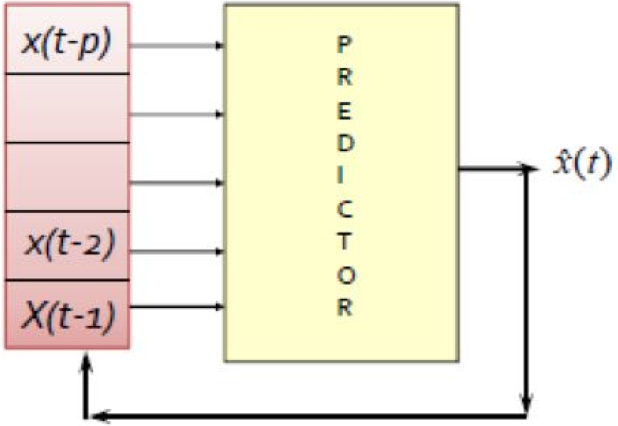
Synthetic Modeling on Multiple Time Series

- Similar to as synthetic modeling on single time series
- But more difficult because we need to capture the relationship among several time series such as the cross-correlation, the spatial burstiness characteristics, etc.
- The most difficulty lies in the relationship between time and space such as when the arrival time series is also modeled together with other time series
- Deep learning techniques such as RNN, LSTM and variations can be applied
- Example: language modeling
- Demo

Predictive Modeling on Single Time Series

- Given a time series, predictive modeling on single time series refers to the process of calculating one of several values ahead, using just the information given by the past values of the time series
- One-step prediction vs. several-step prediction





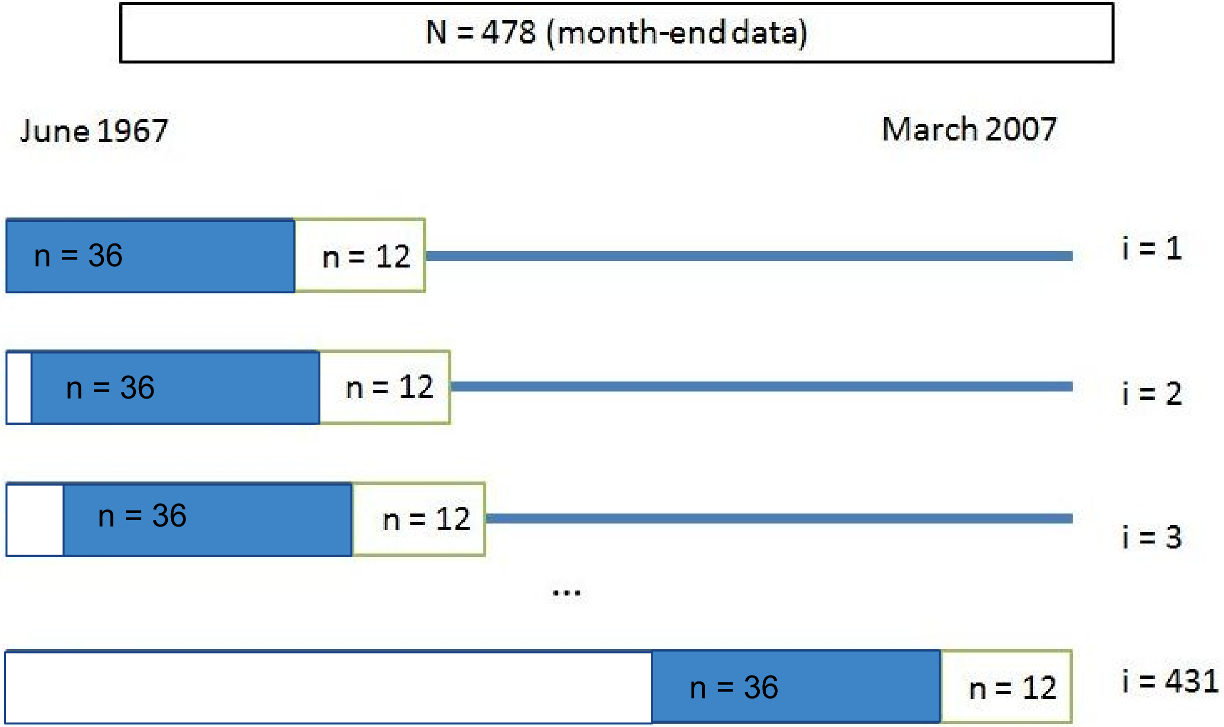
Predictive Modeling on Single Time Series

- Algorithms
  - Deep learning techniques such as RNN, LSTM and variations
  - ARIMA family
  - Holt, Holt-Winter
  - etc.
- Demo

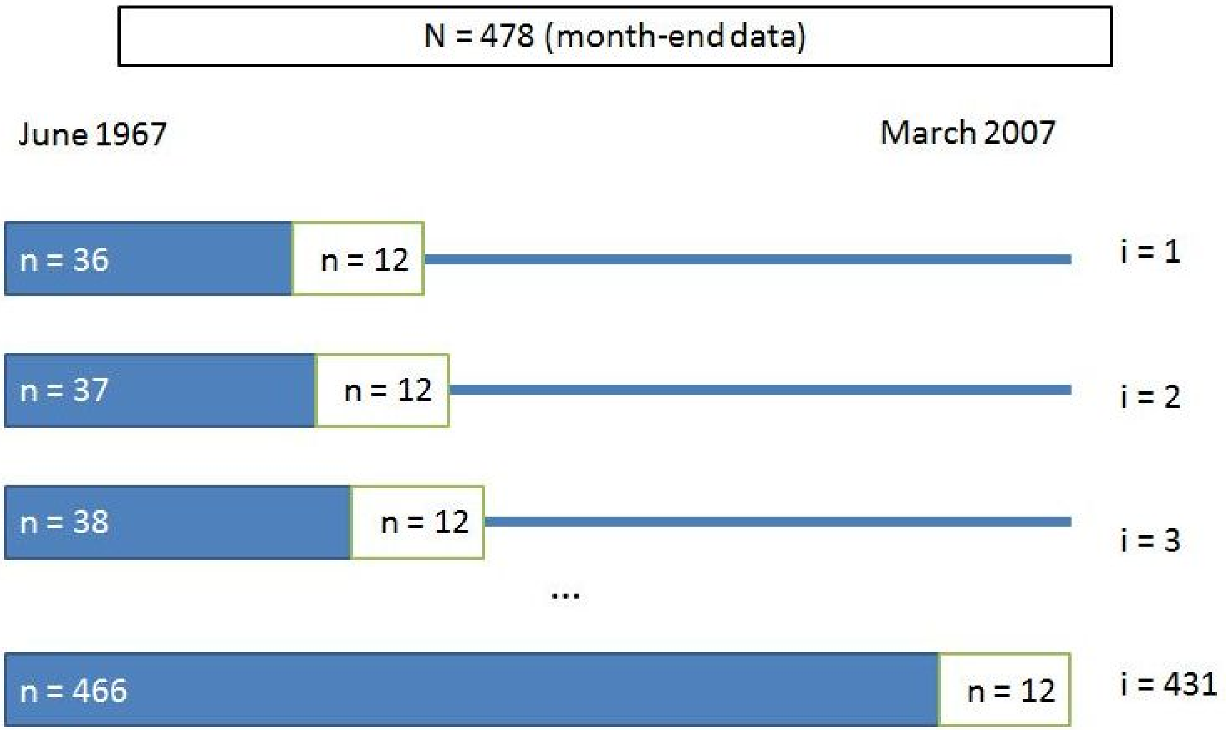
Predictive Modeling on Multiple Time Series

- Similar to “normal” machine learning
- The difference now is that each feature is now a time series
- Need some special techniques for validation
  - Add past days as features and then do cross-validation as normal
  - Use k-fold cross-validation for time series

K-fold Cross-Validation for Time Series



K-fold Cross-Validation for Time Series



Agenda

- Introduction to Time Series
- Time Series Representation
- Time Series Characteristics
- Time Series Modeling & Demo
- Discussion

Discussion (We Learn Together)

- Intuitively, how time series characteristics link us to a machine learning algorithm
  - Long range dependence, i.e. Hurst parameter > 0.7, ⇒ LSTM with 1 time series
  - Short range dependence, i.e. Hurst parameter ~ 0.6 ⇒ RNN with 1 time series
  - Periodicity and dependency ⇒ Periodicity transform + LSTM/RNN with 1 time series
  - Cross-correlation + autocorrelation ⇒ LSTM with multiple time series
  - Temporal locality ⇒ near future prediction needs considered
  - Temporal burstiness ⇒ impact on online learning system
  - etc.

Conclusion

- An overview on time series
- For time series, one usually cares about
  - Overview statistics
  - Distribution
  - Complicated characteristics
- Time series modeling includes 4 problems (and probably more…)
  - 2 modeling problems x 2 modeling types
- Depend on what our problem is, suitable learning algorithms need to be selected for fitting

Q&A

- Thanks for your attendance !