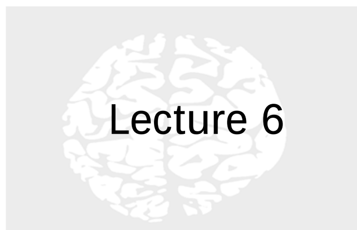
Time Series Problems



Applications

- Prediction
- Noise reduction
- Scientific insight
- Control

Examples

- Weather data
- Climate data
- Tide levels
- Seismic waves
- Sunspots
- Financial markets
- Ecological fluctuations
- EEG data
- ...

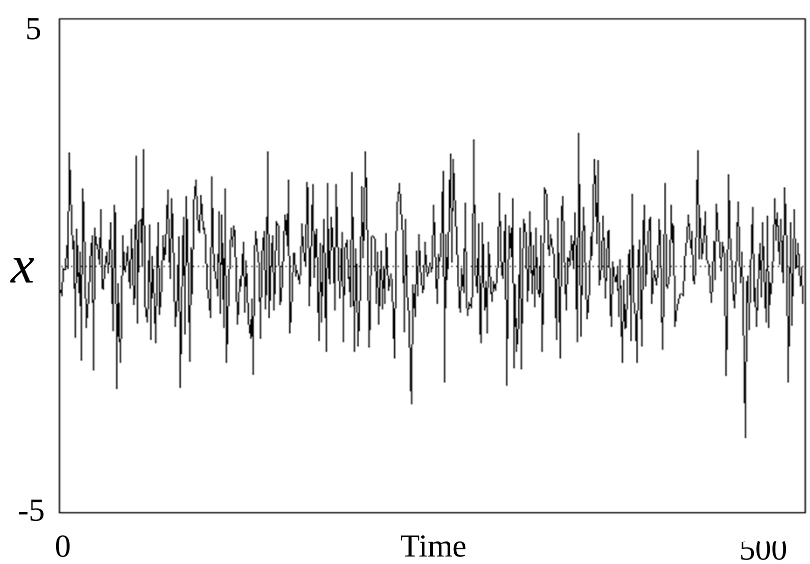
(Non-)Time Series

- Core samples
- Terrain features
- Sequence of letters in written text
- Notes in a musical composition
- Bases in a DNA molecule
- Heartbeat intervals
- ...

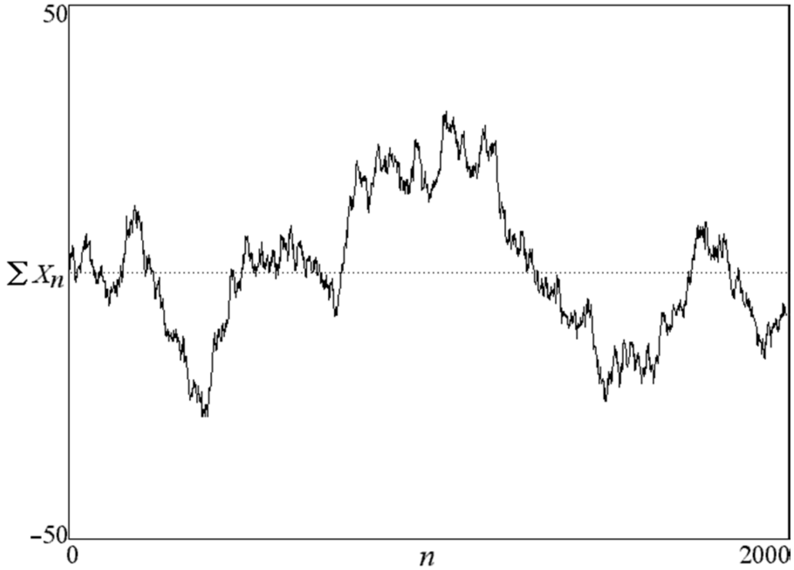
Methods

- Linear (traditional)
  - Fourier Analysis
  - Autocorrelation
  - ARMA
  - ...
- Nonlinear (chaotic)
  - State space reconstruction
  - Correlation dimension
  - Lyapunov exponent
  - Principle component analysis
  - Surrogate data ...

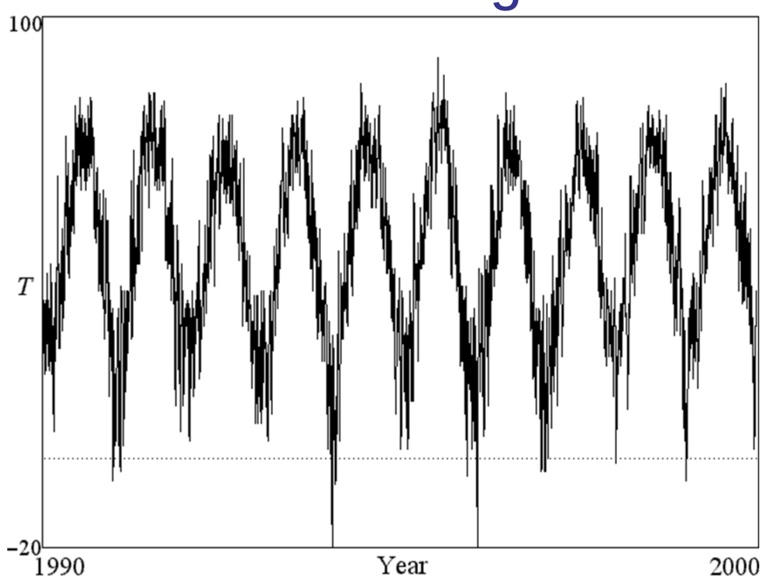
Typical Experimental Data



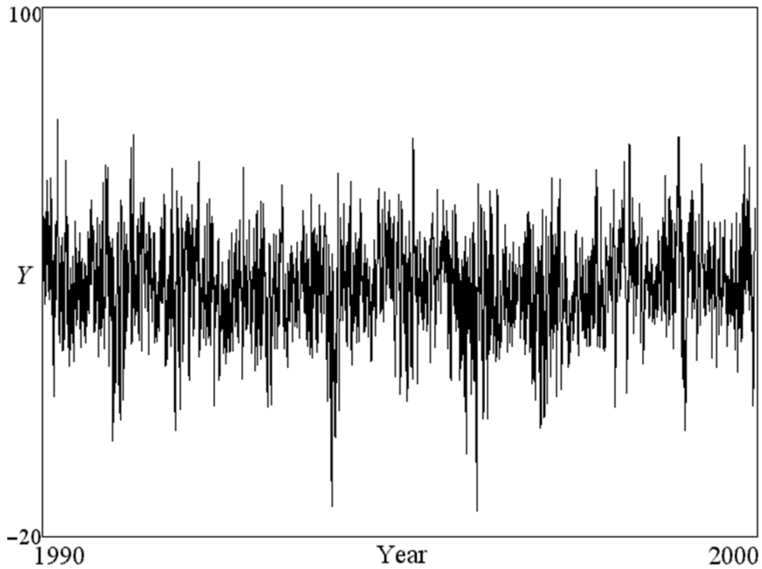
Stationarity



Detrending



Detrended



Practical Considerations

- Calculation speed
- Required number of data points
- Required precision of the data
- Noisy data
- Multivariate data
- Filtered data
- Missing data
- Nonuniformly sampled data
- Nonstationary data

Some General High-Dimensional Models

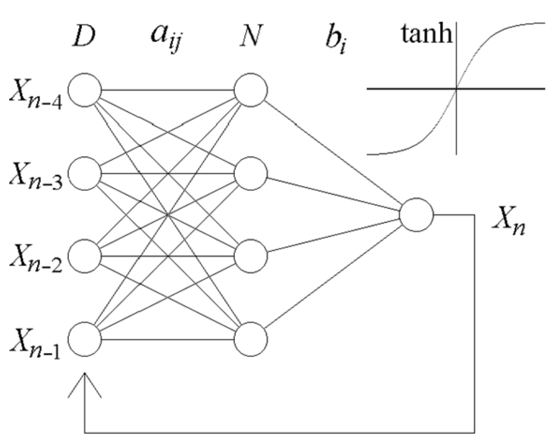
Fourier Series: $x(t)=a_0+\sum_{i=1}^{N} a_i \cos i\omega t + b_i \sin i\omega t$

Linear Autoregression: (ARMA, LPC, MEM…) $x(t)=a_0+\sum_{i=1}^{N} a_i x(t-i)+\text{noise}$

Nonlinear Autogression: (Polynomial Map) $x(t)=a_0+\sum_{i=1}^{N} x(t-i)\left(a_i+\sum_{j=1}^{N} a_{ij} x(t-j)\right)$

Neural Network: $x(t)=b_0+\sum_{i=1}^{N} b_i \tanh \sum_{j=1}^{D} a_{ij} x(t-j)$

Artificial Neural Network



Forecasting

Learning Objectives

- Identify Principles of Forecasting
- Explain the steps in the forecasting process
- Identify types of forecasting methods and their characteristics
- Describe time series and causal forecasting models
- Generate forecasts for different data patterns: level, trend, seasonality, and cyclical
- Describe causal modeling using linear regression
- Compute forecast accuracy
- Explain how forecasting models should be selected

Common Principles of Forecasting

- Many types of forecasting models—differing in complexity and amount of data
- Forecasts are rarely perfect
- Forecasts are more accurate for grouped data than for individual items
- Forecast are more accurate for shorter than longer time periods

Forecasting Steps

- What needs to be forecast?
  - Level of detail, units of analysis & time horizon required
- What data is available to evaluate?
  - Identify needed data & whether it’s available
- Select and test the forecasting model
  - Cost, ease of use & accuracy
- Generate the forecast
- Monitor forecast accuracy over time

Types of Forecasting Models

- Qualitative methods:
  - Forecasts generated subjectively by the forecaster
- Quantitative methods:
  - Forecasts generated through mathematical modeling

Qualitative Methods

Type	Characteristics	Strengths	Weaknesses
Executive opinion	A group of managers meet & come up with a forecast	Good for strategic or new-product forecasting	One person's opinion can dominate the forecast
Market research	Uses surveys & interviews to identify customer preferences	Good determinant of customer preferences	It can be difficult to develop a good questionnaire
Delphi method	Seeks to develop a consensus among a group of experts	Excellent for forecasting long-term product demand, technological changes, and scientific advances	Time consuming to develop

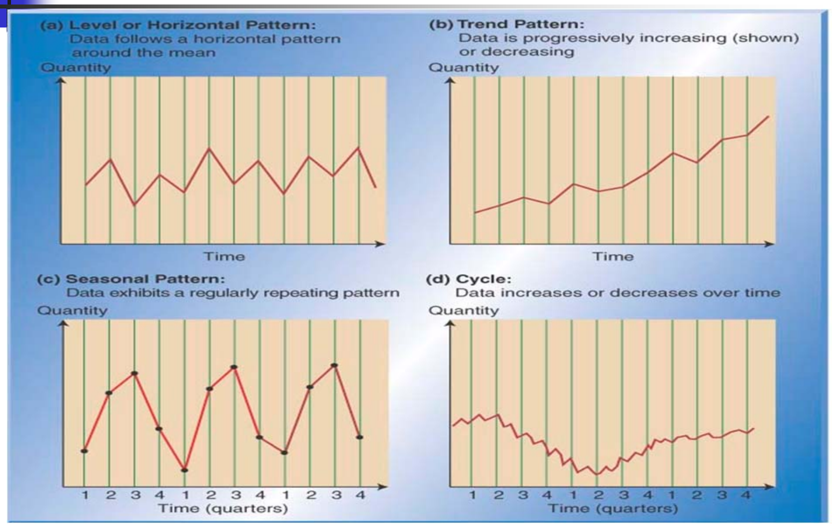
Quantitative Methods

- **Time Series Models:**
  - Assumes the future will follow same patterns as the past
- **Causal Models:**
  - Explores cause-and-effect relationships
  - Uses leading indicators to predict the future
  - E.g. housing starts and appliance sales

Time Series Data Composition

- Data = historic pattern + random variation
- Historic pattern to be forecasted:
  - Level (long-term average)
  - Trend
  - Seasonality
  - Cycle
- Random Variation cannot be predicted

Time Series Patterns



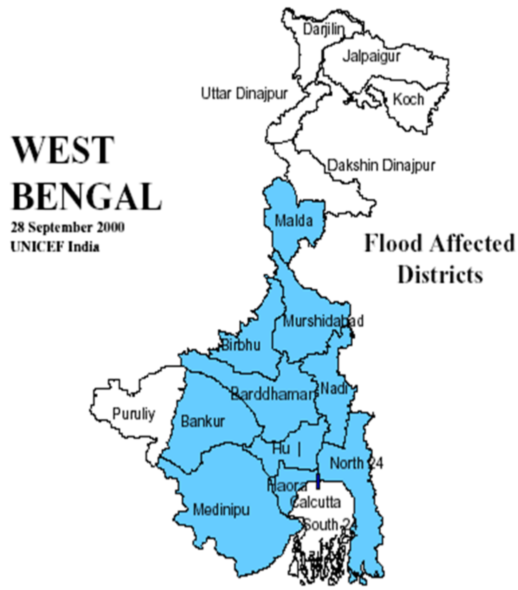
An Example

Application of BP Based Neural Networks to Flood Forecasting

Next Few Minutes

- Introduction
- Conventional Models
- Neural Networks Models and Issues
- Neural Networks as Flood Forecasting Model – Reported Case Study
- Discussion
- Summary

Floods In West Bengal



Definition of Flood Forecast

- The process of estimating the future stages or flows and its time sequence at selected points along the river during floods
- Flood forecasts refer to prediction of the "crest and its time of occurrence" and logical extension to the stages of river above a specified – Warning level (i.e. 1 mt below the danger level)

Utility of Forecast

- Accuracy
- Time of prediction

Flood forecasting service has to be planned around a time-factor keeping in view the following factors

- Availability of operational data (poor, incomplete ...)
- Adoption of appropriate technique
  (accuracy and period of warnings to different locations)
- Dissemination of Forecast

Steps Before Issue of Forecasts

Observation and Collection of data

(Hydrological and Meteorological)

Transmission and communication of data to forecast centers

Analysis of data and formulation of forecast model

Dissemination of forecasts and warnings to authorities of states

Generic Forecast Modeling

Hydro-meteorological, hydrological data

Model

Flood Forecasts

Input Variables

Mapping

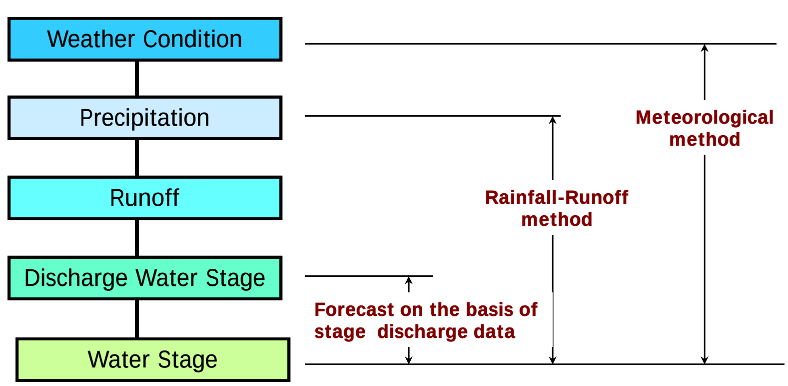
Output

Equation or Mathematical Model

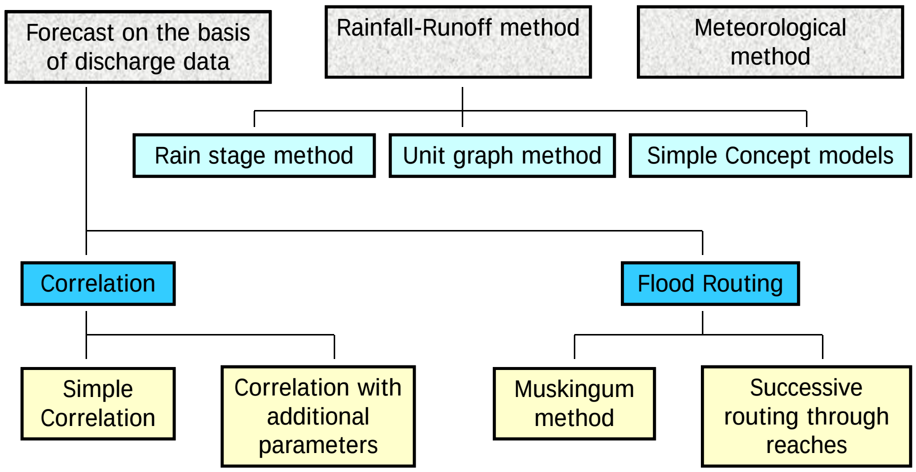
Requirements of Flood Forecasting Model

- Reliable forecast with sufficient warning time (Reasonable degree of accuracy)
- Input data requirement of data model (both for calibration and for operational use) should match the data availability
- Minimum data requirement
- Must have functions which are easy to understand
- Computational procedure simple for field personnel
- Upgradability of the model

Conventional Models



Classification



Is there any other approach to forecast (flood)?

YES

Neural Networks (NN)

NN Forecaster

Hydro-meteorological, hydrological data

↓

Model

↓

Flood Forecasts

Input Variables

Mapping

Output

~~Equation or Mathematical Model~~

Neural Networks

Reported Case Study

Problem Definition

- River flow prediction in segment of Lower Mississippi River

Problem Definition (Contd.)

- Lower Mississippi River (LMR) begins at Cairo, IL at confluence of the Ohio and Upper Mississippi River (UMR) .
- LMR travels downstream around 954 miles
- 1973 saw series of floods in LMR
- Peak flows for the crest stages were over 1.5 million cfs.
- With minimum hydrologic data predict river flow at Memphis, TN and determine the contribution of Ohio river to flooding

Data Collection and Analysis

- Upstream gauge at Thabes, IL on UMR
- Nearby gauge at Metropolis, IL at UMR confluence with Ohio River
- Lateral contribution of tributaries – Obion, Hatchie, Loosahachie, Wolf Rivers in West Tennesse + Rainfall in this river basin

Data Collection and Analysis (Contd.)

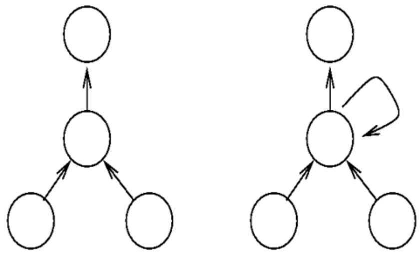
- 16 years (from 1975 to 1990) of daily riverflow from three stations and ten daily rainfall stations (uniformly distributed spatially over the river basin)
- Rainfall-runoff Model
  - Input – Two upstream flows (Ohio and Mississippi Rivers) + total daily rainfall
  - Output – Downstream riverflow
  - First 6 years data for training
  - Next 2 years data for testing
  - Last 8 years data for validation

Neural Networks Models

- **Multilayer perceptrons feed-forward backpropagation architecture, Time Delay, Recurrent** (NeuroSolutions Software)

Recurrent Neural Networks

http://www.idsia.ch/~juergen/rnn.html



Each time a pattern is presented, the unit computes its activation just as in a feed forward network. However its net input now contains a term which reflects the state of the network (the hidden unit activation) before the pattern was seen. When we present subsequent patterns, the hidden and output units' states will be a function of everything the network has seen so far.

The network behavior is based on its history, and so we must think of pattern presentation as it happens in time.

Neural Networks Parameters

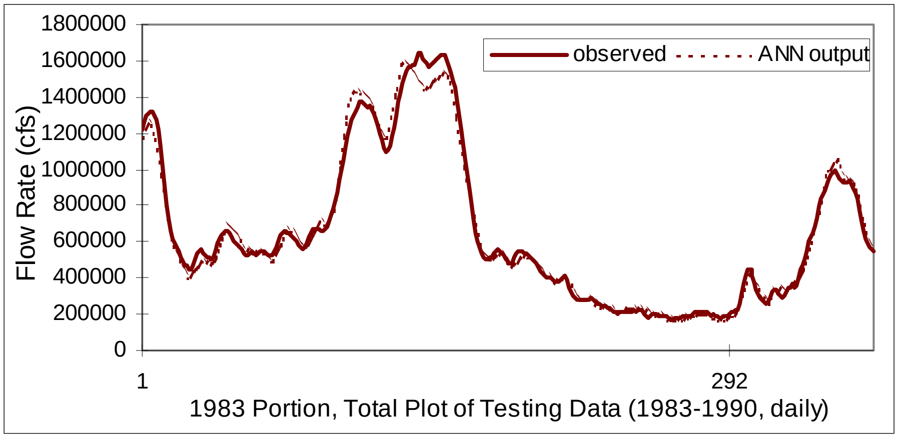
- Number of hidden layers for architecture was ONE
- Performance analysis representation – Mean Square Error (MSE), Normalized Mean Square Error (NMSE), mean/maximum/minimum absolute errors, and correlation coefficient (CC)
- Other NN parameters are missing in the reported paper

NN Performance Evaluation

Scenario 1:

- For Training- Testing-Validation – Fairly accurate results for NMSE and CC
- The CC outcome 0.95,0.93,0.94.
- Spikes matches however, phase existed between observed values and simulated outputs. i.e. consideration of time lags was required

NN Performance Evaluation



NN Performance Evaluation

Scenario 2:

- Study was carried out for 2-day lag for each input series
- Validation overestimated the flow values, but testing set presented excellent results.
- High correlation of downstream gauges was related to 2-day lag riverflow for both upstream gauges

Results Comparison

- Backpropagation
- Backpropagation with time shift input
- Time Delay Neural Networks
- Recurrent Neural Networks

Prediction Reliability of Different Algorithms

		Backpropagation	Backpropagation, time shift input	Time-Delay	Recurrent
Training	NMSE	0.0966	0.0303	0.0194	0.0168
	R	0.9505	0.9847	0.9903	0.9918
Testing	NMSE	0.1448	0.0416	0.0665	0.0652
	R	0.9286	0.9807	0.9679	0.9680
Validation	NMSE	0.1042	0.0344	0.0210	0.0171
	R	0.9475	0.9834	0.9909	0.9922

Failure Case of Forecaster ☺