1. Introduction and Summary

In attempts to understand the world around us, observations are frequently made sequentially over time. Values in the future depend, usually in stochastic manner, on the observations available present. Such dependence makes it worthwhile to predict the future from its past.

Indeed, in this course we will depict underlying dynamics from which the observed data are generated and will therefore forecast and possible control future events.

1.1 Examples of Time Series

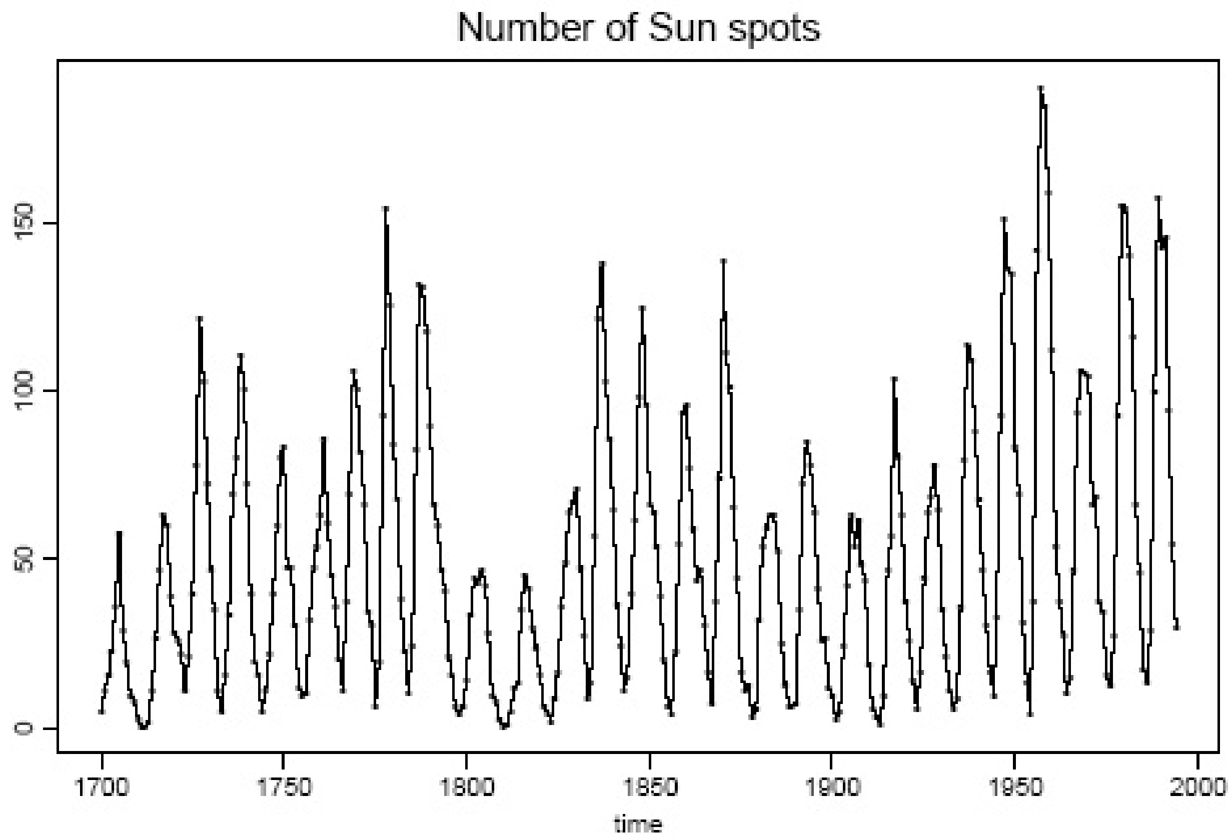
Time Series is a sequence of observations taken sequentially in time. Time series analysis deals with records that are collected over time.

- One distinguishing feature in time series is that the records are usually dependent.
- The background of time series applications is very diverse.
- Depending on different applications, data may be collected hourly, daily, weekly, monthly or yearly, and so on

We use notation such as $\{X_t\}$ or $\{Y_t\}$ $(t = 1, \cdots, T)$ to denote a time series of length T. The unit of the time scale is usually implicit in the notation above.

Example 1.1 (*Sunspot data* from 1770 to 1994) The sunspot numbers from 1770 to 1994 are plotted against time in Figure 1.1. The horizontal axis is the index of time t, and the vertical axis represents the observed value X_t over time t.

Dark spots on the surface of the Sun have consequences in the overall evolution of its magnetic oscillation. They also relate to the motion of the solar dynamo.



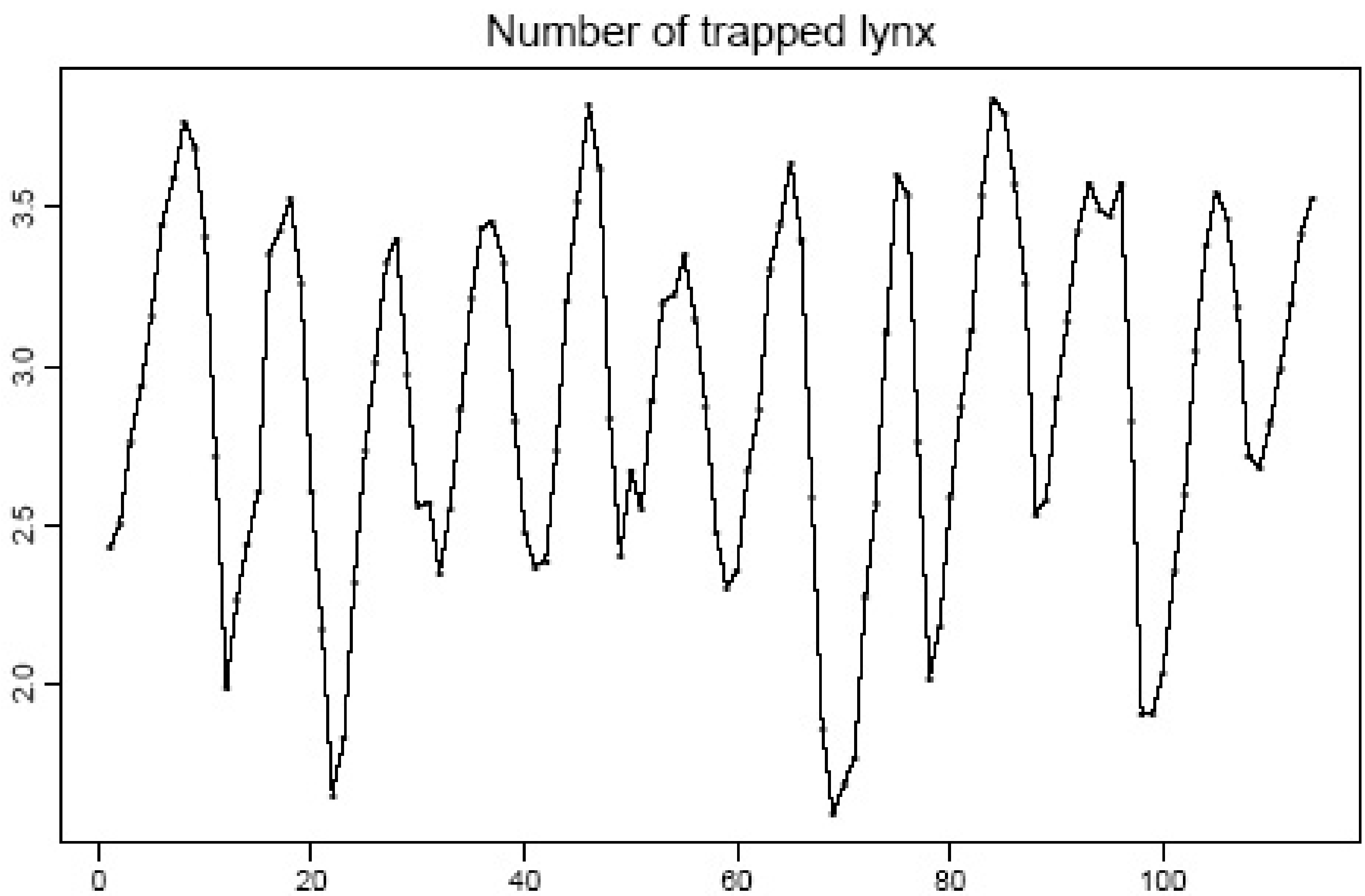
Example 1.2(*Canadian lynx data*) This data set consists of the annual fur returns of lynx at auction in London by the Hudson Bay Company for the period 1821—1934.

It is a proxy of the annual number of the Canada lynx trapped in the Mackenize River district of northwest Canada and reflects to some extent the population size of lynx in the Mackenize River district. Hence it helps us to study the population dynamics of the ecological system in that area.

Figure 1.2 depicts the time series plot of

$$X_t = \log_{10}(\text{number of lynx trapped in year } 1820 + t), t = 1, 2, \cdots, 114.$$

The periodic fluctuation displayed in this time series has profoundly influenced ecological theory. The dataset has been constantly used to examine such concept as “balance-of-nature”, predator and prey interaction, and food web dynamics.


Number of trapped lynx
3.5
3.0
2.5
2.0
0
20
40
60
80
100

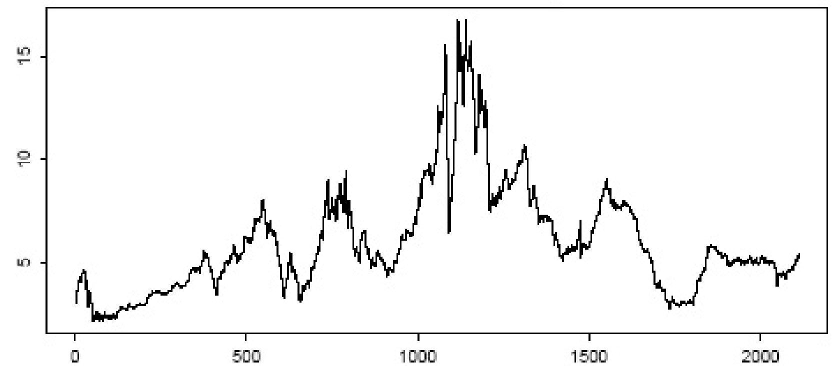
Example 1.3 (*Interest rate data*) Short-term risk-free interest rates play a fundamental role in financial markets. They are directly related to consumer spending, corporate earning, asset pricing, inflation, and the overall economy.

This example concerns the yields of the three-month, six-month, and twelve-month Treasury bills from the secondary market rates (on Fridays). The secondary market rates are annualized using a 360-day year of bank interest and quoted on a discount basis. The data consist of 2,386 weekly observations from July 17, 1959 to September 24, 1999, and are presented in Figure 1.3.

This is a multivariate time series. The correlation matrix among the three series is as follows.

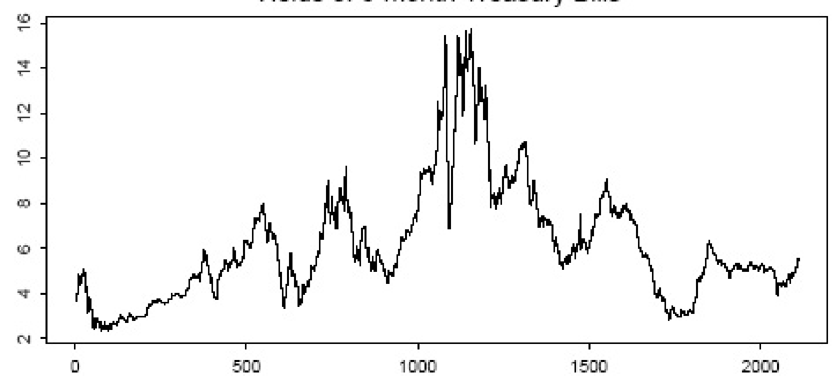
$$\begin{pmatrix} 1.0000 & 0.9966 & 0.9879 \\ 0.9966 & 1.0000 & 0.9962 \\ 0.9879 & 0.9962 & 1.0000 \end{pmatrix}$$

Yields of 3-month Treasury Bills



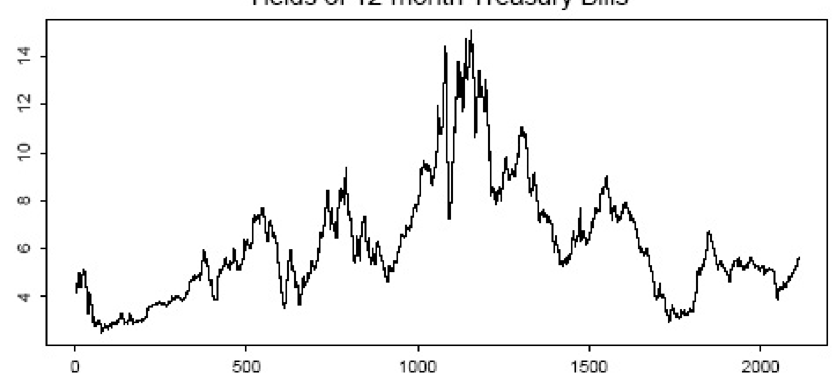
(a)

Yields of 6-month Treasury Bills



(b)

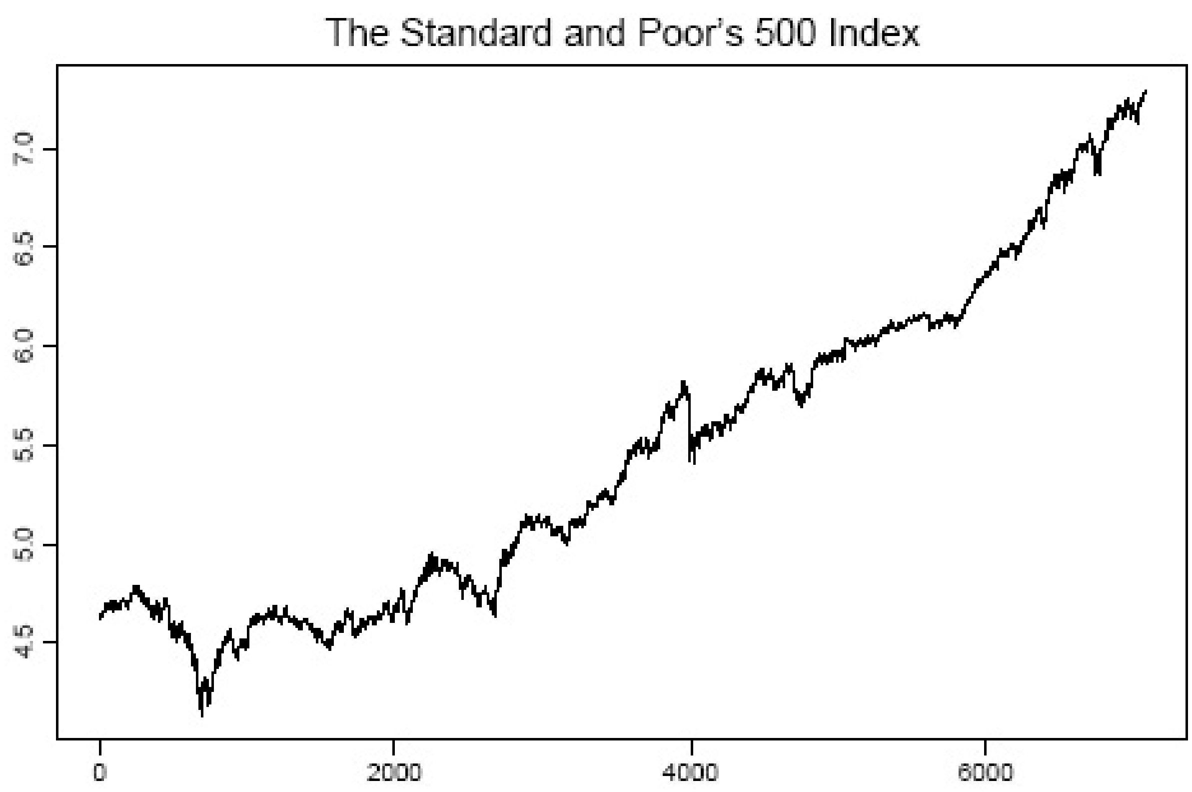
Yields of 12-month Treasury Bills



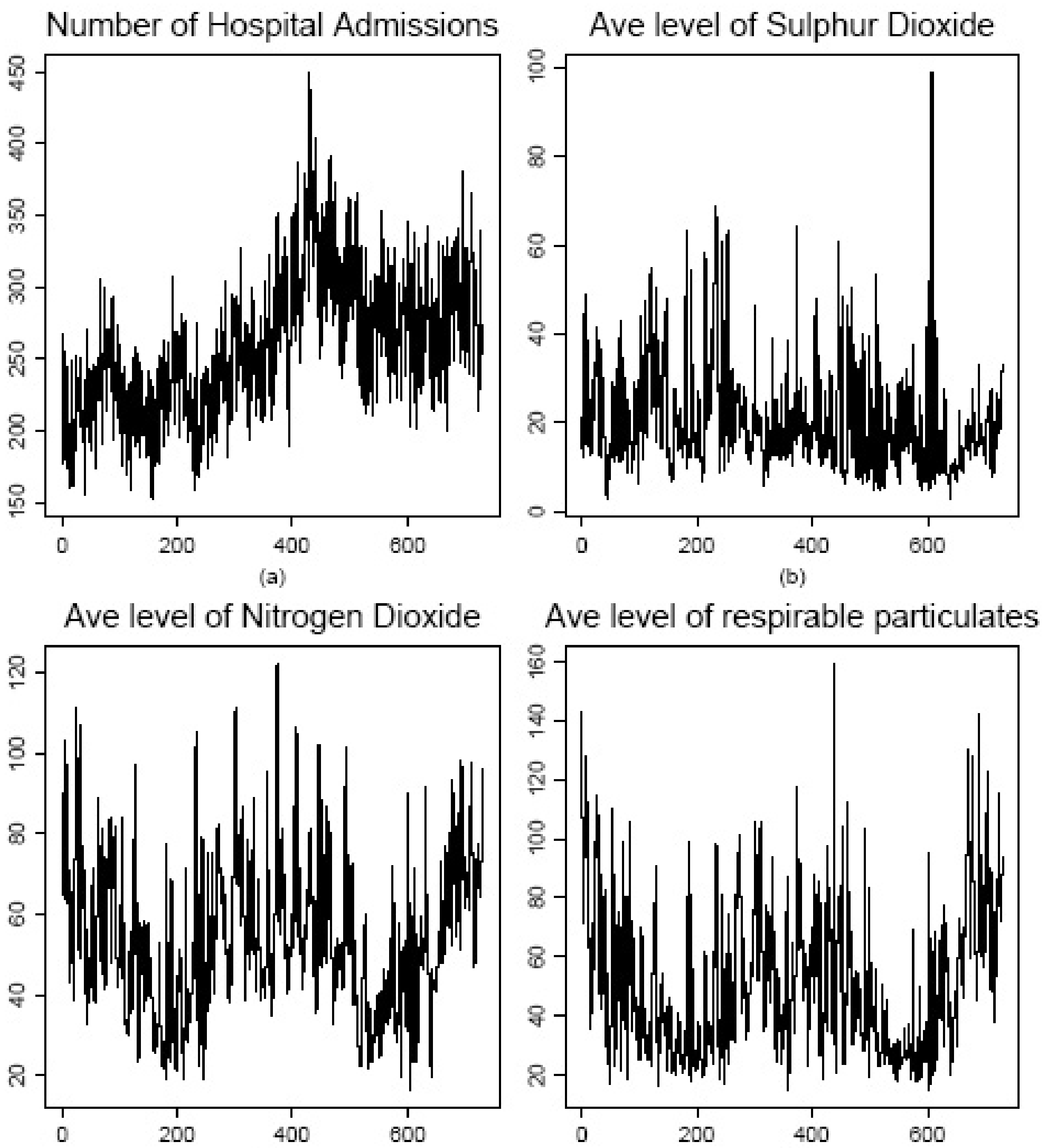
(c)

Example 1.4 (*The Standard and Poor's 500 Index*) The Standard and Poor's 500 index (S&P 500) is a value -weighted index based on the prices of the 500 stocks that account for approximately 70% of the total U.S. equity market capitalization.

Presented in Figure 1.4 are the 7,076 observations of daily closing price of the S&P 500 index from January 3, 1972 to December 31, 1999. The logarithm transform has been applied so that the difference is proportional to the percentage of investment return.

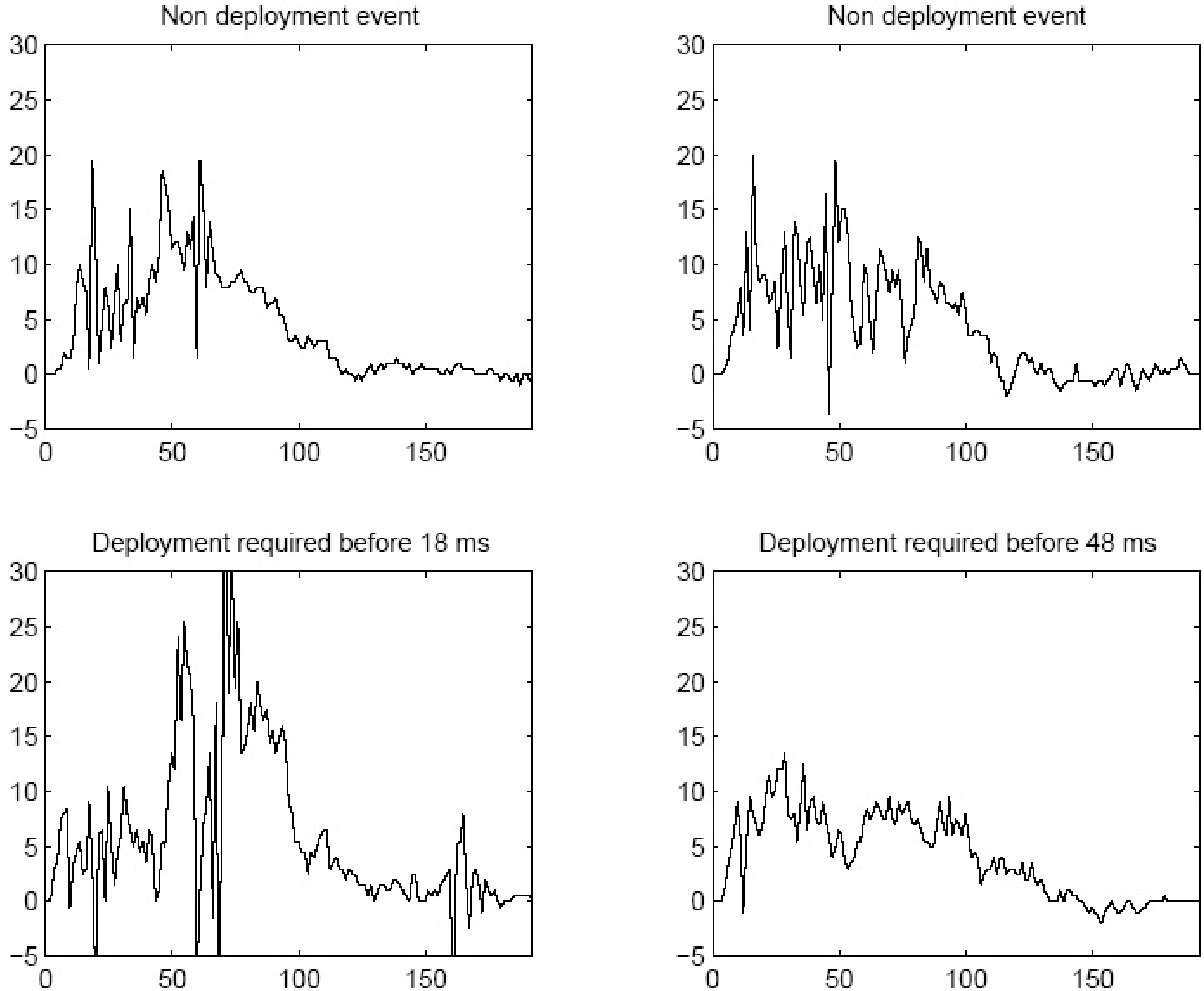


Example 1.5 (*An environment data set*) The environmental condition plays a role in public health. There are many factors that are related to the quality of air that may affect human circulatory and respiratory system. The data set used here comprises daily measurements of pollutants and other environmental factors in Hong Kong between January 1, 1994 and December 31, 1995. We are interested in studying the association between the level of pollutants and other environmental factors and the number of total daily hospital admission for circulatory and respiratory problems.


Number of Hospital Admissions
150
200
250
300
350
400
450
0
200
400
600
(a)
Ave level of Sulphur Dioxide
0
20
40
60
80
100
0
200
400
600
(b)
Ave level of Nitrogen Dioxide
20
40
60
80
100
120
0
200
400
600
Ave level of respirable particulates
20
40
60
80
100
120
140
160
0
200
400
600

Example 1.5 (*Signal processing-deceleration during car crashes*) Time series often appear in signal processing. As an example, we consider the signals from crashes of vehicles. Airbag deployment during a crash is accomplished by a microprocessor-based controller performing an algorithm on the digitized output of an accelerometer. The accelerometer is typically mounted in the passenger compartment of the vehicle. It experiences decelerations of varying magnitude as the vehicle structure collapses during a crash impact.

The observed data in Figure 1.6 are the time series of the acceleration (relative to the driver) of the vehicle, observed at 1.25 milliseconds per sample .



Non deployment event
Non deployment event
Deployment required before 18 ms
Deployment required before 48 ms
30
25
20
15
10
5
0
−5
0
50
100
150

1.2 Objectives of Time Series Analysis

- Unveil the probability law that governs the observed time series.
- Understand the underlying dynamics
- Forecast future events
- Control future events via intervention.

Time series analysis rests on **proper statistical modeling**.

- In selecting a model, interpretability, simplicity, and feasibility play important roles.

- A selected model should reasonably reflect the physical law that governs the data.

- Everything else being equal, a simple model is usually preferable.

- The family of probability models should be reasonably large to include the underlying probability law that has generated the data but should not be so large that defining parameters can no longer be estimated with reasonably good accuracy.

In choosing a probability model,

1. First extracts salient features from the observed data then chooses an appropriate model that possesses such feature.

2. After estimating parameters or functions in the model, verifies whether the model fits the data reasonably well and looks for further improvement whenever possible.

3. Different purposes of the analysis may also dictate the use of different models. For example, a model that provides a good fitting and admits nice interpretation is not necessarily good for forecasting.

1.3 Forecast

Importance of Good Forecast

The ability to form good forecasts has been highly valued throughout history. Since future events involve uncertainly, the forecasts are usually not perfect. The objective of forecasting is to reduce the forecast error: to produce forecast that are seldom incorrect and that have small forecast errors. In business, industry, and government, policymakers must anticipate their decisions depend on forecast, and they expect these forecasts to be accurate.

Classification of Forecast methods

- Qualitative technique
  Qualitative or subjective forecast methods are intuitive, largely educated guesses that may or may not depend on past data. Usually these forecasts cannot be reproduced by someone else.

- Quantitative technique
  Forecasts that are based on mathematical or statistical models are called quantitative.
  Deterministic model

$$Y = f(X_1, \ldots, X_p; \beta_1, \ldots, \beta_m)$$

  Probabilistic or stochastic models

$$Y = f(X_1, \ldots, X_p; \beta_1, \ldots, \beta_m) + \text{noise}$$

  Frequently the function form f and the coefficient are not known and have to be determined from past data. Usually the data occur in time-ordered sequences referred to as ***time series***.

Conceptual Framework of a Forecast System In general, a quantitative forecast system consists of two major components, as illustrated in the following figure

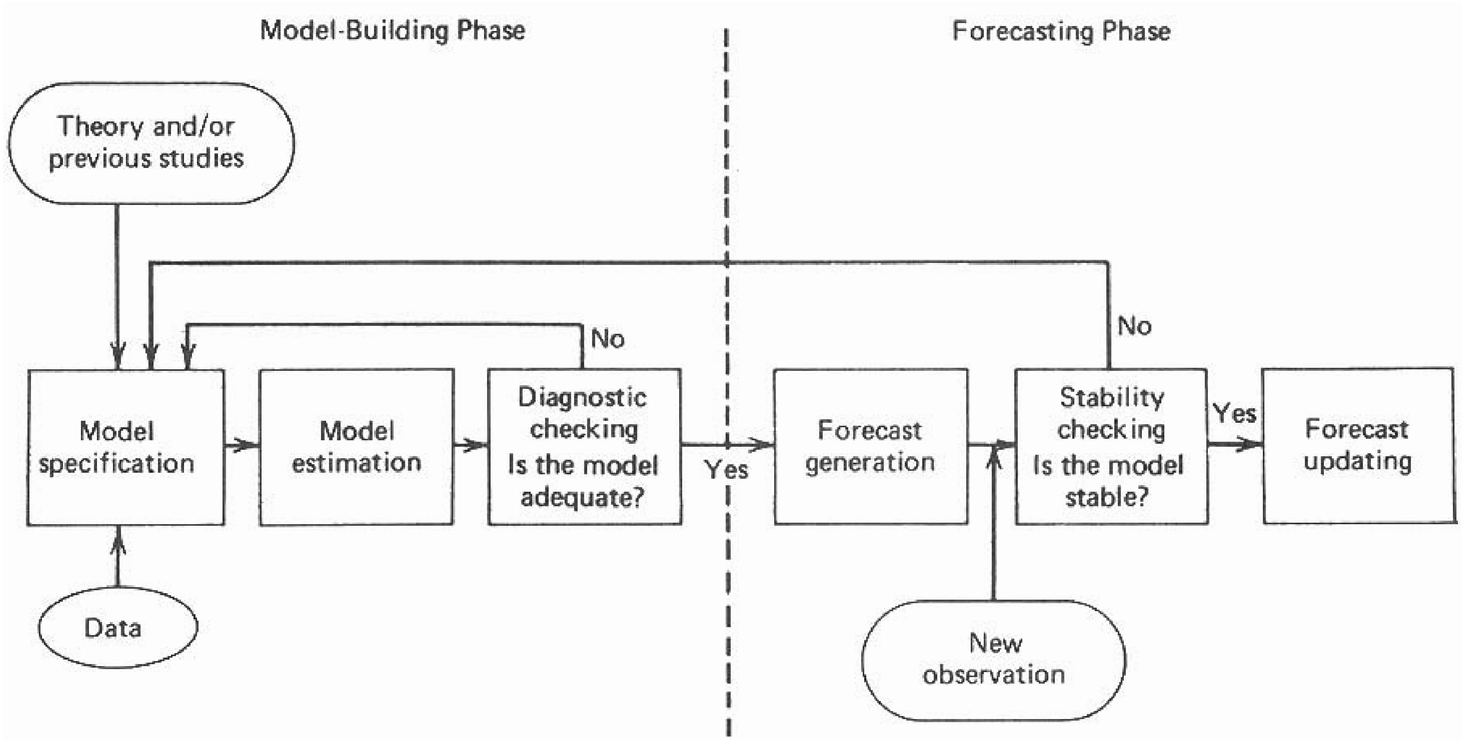


Figure 1.1. Conceptual framework of a forecasting system.

- At the first stage, the model-building-stage, a forecast model is constructed from pertinent data and available theory. The tentatively entertained model usually contains unknown parameters ; an estimation approach, such as least squares , can be used to determine these constants. Finally, the forecaster must check the adequacy of the fitted model. If the model is unsatisfactory, it has to be respecified, and the iterative cycle of model specification-estimation-diagnostic checking must be repeated until a satisfactory model is found.

- At second stage, the forecasting stage, the final model is used to obtain the forecasts. Since these forecast depend on the specified model, one has to make sure that the model and its parameters stay constant during the forecast period. The stability of the forecast model can be assessed by checking the forecasts against the new observations. Forecast error can be calculated, and possible changes in model can be detected.

Choice of A particular forecast model Among many other forecast criteria, the choice of the forecast model or technique depends on

1. what degree of accuracy is required,
2. what the forecast horizon is,
3. how high a cost for producing the forecasts can be tolerated,
4. what degree of complexity is required,
5. what data are available.

Forecast Criteria

The most important criterion for choosing a forecast method is its accuracy, or how closely the forecast predicts the actual event.

- Actual observation at time t with z_t, and its forecast, which uses the information up to and including $t-1$, with $z_{t-1}(1)$
- Future forecast error $z_t - z_{t-1}(1)$
- Since z_t has not been observed, its value is unknown; we can talk only about its expected value, conditional on the observed history up to and including time $t-1$. Hence we could define
  - The mean absolute error $\mathrm{E}|z_t - z_{t-1}(1)|$
  - The mean square error $\mathrm{E}[z_t - z_{t-1}(1)]^2$
  - The forecast that minimize the mean square error are called ***minimum mean square error (MMSE) forecast***

1.4 Linear Time Series Models

White Noise Processes

A stochastic process $\{X_t\}$ is called *white noise*, denote $\{X_t\} \sim WN(0, \sigma^2)$, if

$$\mathrm{E}X_t = 0, \mathrm{Var}(X_t) = 0, \text{and } \mathrm{Cov}(X_i, X_j) = 0, \text{for all } i \neq j.$$

- White noise is defined by the properties of its first two moments only
- $Gaussian\ Process$ When all of the finite dimensional distributions are Gaussian (normal)

AR Models

An autoregressive model of order $p \geq 1$ is defined as

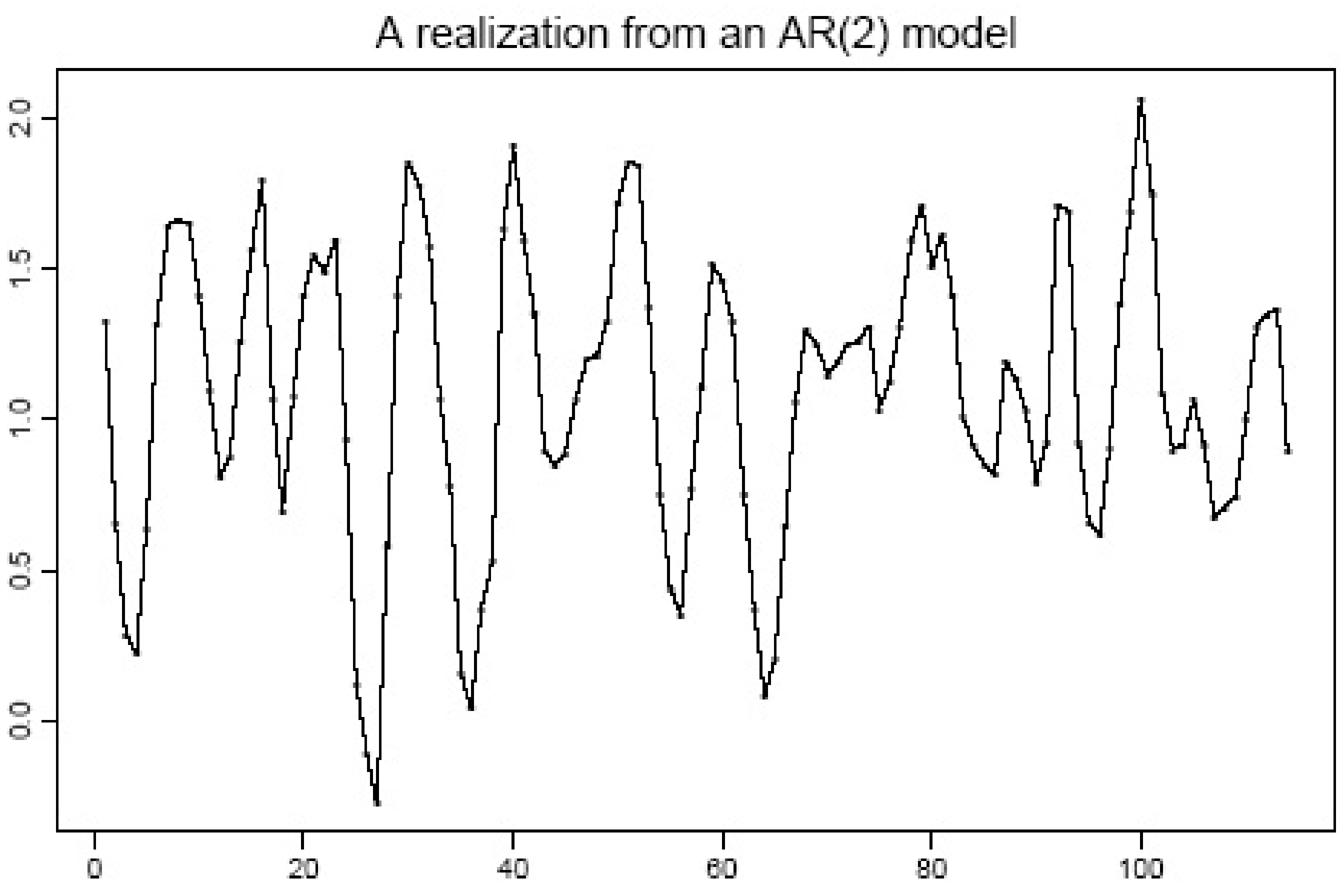
$$X_t = b_1 X_{t-1} + \cdots + b_p X_{t-p} + \varepsilon_t,$$

where $\{\varepsilon_t\} \sim \mathsf{WN}(0, \sigma^2)$. This model is denoted by $\{X_t\} \sim \mathsf{AR}(p)$.

The following Figure shows a realization of a time series of length 114 from an AR(2)-model.

$$X_t = 1.07 + 1.35X_{t-1} - 0.72X_{t-2} + \varepsilon_t$$

with $\{\varepsilon_t\} \sim_{\text{i.i.d.}} \mathcal{N}(0, 0.24^2)$.



MA Models

A moving average process with order $q \geq 1$ is defined as

$$X_t = \varepsilon_t + a_1\varepsilon_{t-1} + \cdots + a_q\varepsilon_{t-q},$$

where $\{\varepsilon_t\} \sim \mathsf{WN}(0, \sigma^2)$. We write $\{X_t\} \sim \mathsf{MA}(q)$.

- The implementation of an MA-model is more difficult than that of an AR-model, since $\{\varepsilon_t\}$ is unobservable.
- They provide parsimonious representations of time series exhibiting MA-like correlation structure.
- MA models lies in their theoretical tractability

ARMA Models

An autoregressive moving average (ARMA) model defined as

$$X_t = b_1 X_{t-1} + \cdots + b_p X_{t-p} + \varepsilon_t + a_1 \varepsilon_{t-1} + \cdots + a_q \varepsilon_{t-q},$$

where $\{\varepsilon_t\} \sim \mathsf{WN}(0, \sigma^2), p, q \geq 0$. We write $\{X_t\} \sim \mathsf{ARMA}(p, q)$.

ARMA models are one of the most frequently used families of parametric models in time series analysis. This is due to their flexibility in approximating many stationary processes.

ARIMA Models

A time series $\{Y_t\}$ is called an autoregressive integrated moving average (ARIMA) process with order p, d and q, if its d-order difference

$$X_t = (1 - B)^d Y_t$$

is a stationary ARMA(p, q) process, where $d > 1$ is an integer, and B denotes backshift operator, which is defined as $B^k X_t = X_{t-k}, k = \pm 1, \pm 2, \cdots$, namely,

$$b(B)(1 - B)^d Y_t = a(B)\varepsilon_t,$$

and

$$b(z) = 1 - b_1 z - \cdots - b_p z^p, \quad a(z) = 1 + a_1 z + \cdots + a_q z^q.$$

As an illustration, we have simulated a time series of length 200 from the ARIMA(1,1,1) model

$$(1 - 0.5B)(1 - B)Y_t = (1 + 0.3B)\varepsilon_t, \{\varepsilon_t\} \sim_{\text{i.i.d}} \mathcal{N}(0, 1).$$