CHAPTER 1

Introduction

"Don't never prophesy: If you prophesies right, ain't nobody going to remember and if you prophesies wrong, ain't nobody going to let you forget."

– Mark Twain

Good forecasts are vital in many areas of scientific, industrial, commercial and economic activity. This book is concerned with *time-series forecasting*, where forecasts are made on the basis of data comprising one or more time series. A *time-series* is a collection of observations made sequentially through time. Examples include (i) sales of a particular product in successive months, (ii) the temperature at a particular location at noon on successive days, and (iii) electricity consumption in a particular area for successive one-hour periods. An example is plotted in Figure 1.1.

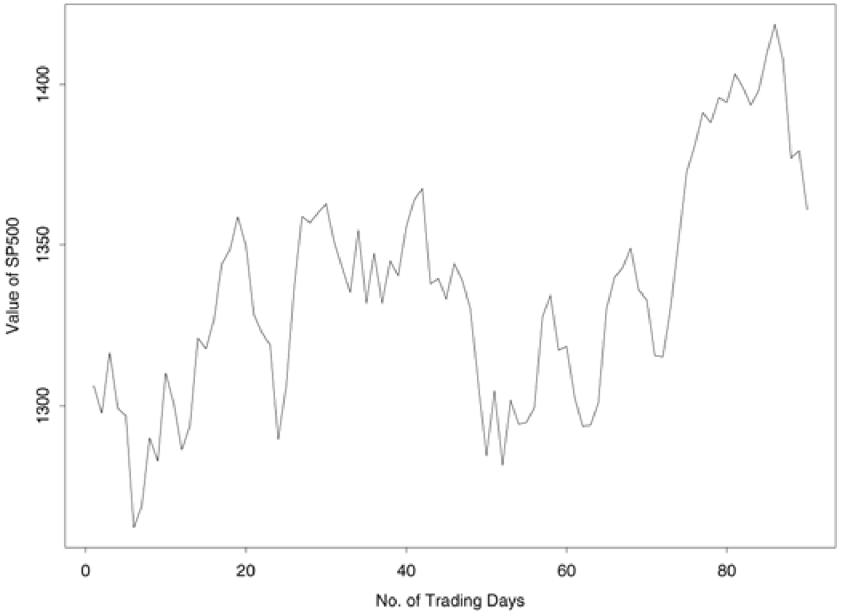

Figure 1.1. *A graph showing the Standard & Poor (S & P) 500 index for the U.S. stock market for 90 trading days starting on March 16 1999. (Note that values for successive trading days are plotted at equal intervals even when weekends or public holidays intervene.)*

Applications of time-series forecasting include:

1. Economic planning
2. Sales forecasting
3. Inventory (or stock) control
4. Production and capacity planning
5. The evaluation of alternative economic strategies
6. Budgeting
7. Financial risk management
8. Model evaluation

Most of the applications on the above list are self-explanatory. For example, good forecasts of future sales will obviously make it easier to plan production. However, the reader may not have realized that forecasts can help in *model evaluation*, when trying to fit time-series models. Checks on fitted models are usually made by examining the goodness-of-fit of the same data used to estimate the model parameters (called the *in-sample* fit). However, time-series data provide an excellent opportunity to look at what is called *out-of-sample* behaviour. A time-series model will provide forecasts of new future observations which can be checked against what is actually observed. If there is good agreement, it will be argued that this provides a more convincing verification of the model than in-sample fit; – see Sections 3.5.3 and 8.5.4.

The book does not attempt to cover the specialist areas of *population* and *weather forecasting*, although time-series techniques could, for example, be used to forecast a specific variable such as temperature. We also exclude specific mention of *probability forecasting*, where the aim is to predict the probability of occurrence of a specific event outcome such as a turning point, a change in interest rate or a strike.[1]

This book is not intended to be a comprehensive manual of forecasting practice, as several books are already available which describe in detail how to apply particular forecasting methods. Rather this book aims to give an overview of the many methods available, and to provide appropriate up-to-date references for the reader who wants more detail. There is emphasis on the intuitive ideas underlying the different methods, some comments on recent research results, and some guidance on coping with practical problems. The inter-relationships between the different methods are explored and the methods are compared from both theoretical and empirical points of view. After revising the basics of time-series analysis in Chapter 2, Chapter 3 discusses model building for a single series, both for particular classes of model and in general terms. Chapters 4 and 5 look at univariate and multivariate forecasting methods, respectively, while Chapter 6 discusses the evaluation of forecasts and gives advice on

[1] See, for example, the special issue on Probability Forecasting, *Int. J. of Forecasting*, 1995, No. 1.

which method to choose in different contexts. Chapters 7 and 8 cover two additional important topics, namely different ways of calculating prediction intervals (rather than point forecasts), and the effect of model uncertainty on forecast accuracy, especially in regard to the tendency for empirical results to suggest that prediction intervals are generally too narrow.

This opening introductory chapter begins by discussing the different categories of forecasting method in Section 1.1, while some preliminary practical questions are raised in Section 1.2. Section 1.3 contains a 'Public Health Warning' on the dangers of extrapolation, and Section 1.4 explains the important distinction between forecasts that are made 'in-sample' (and which are therefore not genuine forecasts) and those which are 'out-of-sample'. The chapter ends in Section 1.5 with a brief overview of relevant literature.

1.1 Types of forecasting method

Suppose we have an observed time series $x_1, x_2, \ldots, x_N$ and wish to forecast future values such as x_{N+h}. The integer h is called the *lead time* or the *forecasting horizon* (h for horizon) and the forecast of x_{N+h} made at time N for h steps ahead will be denoted by $\hat{x}_N(h)$. Note that it is essential to specify both the time the forecast is made *and* the lead time. Some of the literature does not do this and instead uses an ambiguous notation such as $\hat{x}_{N+h}$ for forecasts of X_{N+h} regardless of when the forecast was made.

A *forecasting method* is a procedure for computing forecasts from present and past values. As such it may simply be an algorithmic rule and need not depend on an underlying probability model. Alternatively it may arise from identifying a particular model for the given data and finding optimal forecasts conditional on that model. Thus the two terms 'method' and 'model' should be kept clearly distinct. It is unfortunate that the term 'forecasting model' is used rather loosely in the literature and is sometimes wrongly used to describe a forecasting *method*.

Forecasting methods may be broadly classified into three types:

(a) *Judgemental forecasts* based on subjective judgement, intuition, 'inside' commercial knowledge, and any other relevant information.

(b) *Univariate methods* where forecasts depend only on present and past values of the single series being forecasted, possibly augmented by a function of time such as a linear trend.

(c) *Multivariate methods* where forecasts of a given variable depend, at least partly, on values of one or more additional time series variables, called predictor or explanatory variables. Multivariate forecasts may depend on a multivariate model involving more than one equation if the variables are jointly dependent.

More generally a forecasting method could combine more than one of the above approaches, as, for example, when univariate or multivariate forecasts are adjusted subjectively to take account of external information which is difficult to express formally in a mathematical model.

This book focuses on univariate and multivariate time-series methods, and does not attempt to cover judgemental forecasting. Instead the reader is referred, for example, to the literature review by Webby and O'Connor (1996). The most famous judgmental method is probably that called the *Delphi technique* (e.g. Rowe and Wright, 1999), which aims to find a consensus of opinion for a group of 'experts', based on a series of questionnaires filled in by individuals and interspersed with the controlled feedback of opinion and information from other experts. Empirical results suggest that judgemental methods sometimes work well and sometimes do not (as is true for all forecasting methods!). However, statistical methods tend to be superior in general, provided that the latter are a practical proposition. Of course, there are occasions when model-based methods are not practical, perhaps because some essential information is not available, and then judgmental methods have to be used anyway. In any case, it is not sensible to pretend that the modelling and judgemental approaches are completely distinct, as it often helps to combine the two approaches and get the best of both worlds. In particular, many macroeconomic forecasts are obtained by making adjustments to model-based forecasts, perhaps by adding or subtracting an appropriate constant (sometimes called an intercept correction). It is, however, unfortunate that it is not always made clear how such adjustments are made in order to arrive at a final forecast. As might be expected, the integration of a judgemental and statistical approach can improve accuracy when the analyst has good domain knowledge but can harm accuracy when judgement is biased or unstructured (Armstrong and Collopy, 1998).

While the techniques commonly known as 'judgemental forecasting' are not covered in this book, it should be clearly understood that some element of judgement *is* always involved in forecasting, even when using what is normally regarded as an 'objective' statistical method. Good time-series modelling, like all statistical model building, involves the use of sound subjective judgement in assessing data, selecting a model, and interpreting the results. This use of judgement *will* be covered in what follows.

An alternative important way of classifying forecasting methods is between *automatic* methods, which require no human intervention, and *non-automatic* methods, which do. As one example of this important distinction, it is instructive to compare inventory control with economic planning. In inventory control there may be hundreds, or even thousands, of items to monitor. It is quite impossible to fit separate models to each individual time series of sales. Rather, a simple, automatic method is normally used for the whole range of items. In contrast, economic planning requires the analyst to carefully build an appropriate model describing the relationship between relevant economic variables, after which 'optimal' forecasts can be produced from the model. This method is certainly *not* automatic.

Yet another classification is between *simple* and *complicated* methods. Univariate methods are generally simpler than multivariate methods and it is always a matter of judgement as to how much effort should go into making a prediction. For some economic series, it is hard to beat the very simple forecast of using the latest observation to predict the next one. A deterministic series is also easy to forecast when the model is known. For example, a series which changes periodically through time in a systematic way is easy to forecast once a full cycle has been observed. At the other extreme, a series of independent observations is also 'easy' (or impossible) to predict as you cannot do better than use the overall mean value. Our main interest is in series which are in-between these extremes. They cannot be predicted exactly but they do contain structure which can be exploited to make better forecasts.

1.2 Some preliminary questions

In forecasting, as in any statistical exercise, it is essential to carry out any necessary preliminary work. In particular, it is important to *formulate the problem* carefully (see Chatfield, 1995a, Chapter 3). The analyst must (i) *ask questions* so as to get sufficient background information, (ii) *clarify the objectives* in producing forecasts, and (iii) find out exactly how the forecast will be used. The *context* is crucial in all of this. Ideally forecasts should be an integral part of the planning system and not a separate exercise. This is sometimes called a *systems approach*. It requires that the statistician talks to the people who will actually *use* the forecasts. A relatively simple forecasting method, which is widely understood, may be preferred.

A key related question is whether the forecasts actually influence the outcome. In some situations the forecast is used as a *target* value, while in others the forecasts are used to suggest control action. For example, a sales forecast may become a target in that workers will try to achieve sales of at least the forecast value, while a forecast of an increasing death rate for a particular disease may lead to preventive action to try to reduce the spread of the disease. Forecasts which prompt control action will be self-defeating, and yet such forecasts can be very useful even though they may not score well in terms of accuracy. In one sense, short-term weather forecasting is easier than econometric forecasting, as short-term weather forecasts cannot influence the weather, whereas economic forecasts may influence governmental policy. However, note that long-term weather forecasting (over a period of years, rather than days) could affect the outcome, in that predictions of global warming, for example, may influence government policy to try to reduce greenhouse gases, so as to prevent unwanted changes to weather patterns.

The analyst should also find out if forecasts only are required, or if there is a need for a descriptive, interpretable *model* and whether the forecasts are going to be used for control purposes. As always there is no point in giving the RIGHT answer to the WRONG question (called an *error of the third kind!*). For example, a model that gives a good *fit* to a set of past data may or may not be the most useful model for predicting future values. Fitting past values and forecasting future values are two quite different applications of a model. Similarly the best-fit model may or may not be the most helpful model for providing 'policy advice', such as deciding whether a proposed change in tax rates is likely to result in a beneficial effect on future unemployment. The sort of problem that arises in practice is that economic theory may suggest that one particular variable should be an important explanatory variable in an econometric model but, in a given set of recent data, it is found that the variable has been held more or less constant over the relevant period. Then time-series methods are likely to show that the inclusion or exclusion of the variable makes little difference to the overall fit of a model in regard to explaining the variation in the dependent (or response) variable. Furthermore, on the basis of the given data, it may be impossible to assess the effect of changes in the explanatory variable on forecasts of the response variable. Thus a model derived empirically will be useless for predicting changes in the response variable if a substantial change to the explanatory variable is envisaged, even though the analyst knows that the variable cannot be disregarded. Thus an appropriate model will need to be constructed using economic theory rather than (just) goodness-of-fit to past data.

The other main preliminary task is to make a careful assessment of the available data (sometimes called the *information set*). Have the 'right' variables been recorded to the 'right' accuracy? How much of the data is useful and usable? Are there obvious errors, outliers and missing observations? There is little point in putting a lot of effort into producing forecasts if the data are of poor quality in the first place.

The analyst may also have to ask various additional questions such as how many series are there to forecast, how far ahead are forecasts required, and what accuracy is desired in the forecasts (it may be wise to dampen unrealistic expectations). The forecaster must also decide whether to compute a point forecast, expressed as a single number, or an interval forecast. The latter is often more desirable, though rather neglected in the literature.

The answers to all the above questions determine, in part, which forecasting method should be chosen, as do more pragmatic considerations such as the skill and experience of the analyst and the computer software available. The analyst is advised to use a method he or she feels 'happy' with and if necessary to try more than one method. A detailed comparison of the many different forecasting methods is given in Chapter 6, based on theoretical criteria, on empirical evidence, and on practical considerations.

The importance of clarifying objectives cannot be overstressed. The forecasting literature concentrates on *techniques* – how to implement particular forecasting methods – whereas most forecasters probably need much more help with the *strategy* of forecasting. For example, there is a plenty of software available to make it easy to fit a class of time-series models called ARIMA models (see Section 3.1) but it is still hard to know *when* to use an ARIMA model and *how* to choose which ARIMA model to use.

1.3 The dangers of extrapolation

It is advisable to include here a brief warning on the dangers of forecasting. Time-series forecasting is essentially a form of extrapolation in that it involves fitting a model to a set of data and then using that model outside the range of data to which it has been fitted. Extrapolation is rightly regarded with disfavour in other statistical areas, such as regression analysis. However, when forecasting the future of a time series, extrapolation is unavoidable. Thus the reader should always keep in mind that forecasts generally depend on the future being like the past.

Forecasts also depend on the assumptions which are (implicitly or explicitly) built into the model that is used, or into the subjective judgements that are made. Thus forecasts are generally *conditional* statements of the form that "if such-and-such behaviour continues in the future, then ... ". It follows that one should always be prepared to modify forecasts in the light of any additional information, or to produce a range of different forecasts (rather than just one forecast) each of which is based on a known set of clearly stated assumptions. The latter is sometimes called *scenario forecasting* – see Section 8.5.2 – and Schumacher (1974) is correct when he says that "long-term feasibility studies, based on clearly stated assumptions, are well worth doing". However, Schumacher also says that long-term forecasts "are presumptious", but it is not clear to me when a "feasibility study" becomes a forecast. What we can say is that any really long-term forecast is liable to be way off target. Recent examples I have seen, which I expect to be wrong, include traffic flow in 2020 and world population in 2050! Fortunately for the forecaster, most people will have forgotten what the forecast was by the time it 'matures'. Celebrated examples, which have not been forgotten, include the founder of IBM predicting "a world market for about five computers" in 1947, and the President of Digital Equipment predicting that " there is no reason for any individual to have a computer in their home" in 1977. Going further back into history, I like the quote attributed to U.S. President Hayes in 1876, after he had witnessed a telephone call, that it was "an amazing invention but who would ever want to use one?"

Of course, forecasts can even go horribly wrong in the short-term when there is a sudden change or 'structural break' in the data – see Section 8.5.5. One famous example is that made by a Professor of Economics at Yale University in September 1929, when he said that "Stock prices have reached what looks like a permanently high plateau". This was just before the stock market 'crash', which led on to the Depression!

1.4 Are forecasts genuinely out-of-sample?

"*Prediction is very difficult, especially if it's about the future*" – Nils Bohr

In a real forecasting situation, the analyst typically has data up to time N, and makes forecasts about the future by fitting a model to the data up to time N and using the model to make projections. If $\hat{x}_N(h)$ only uses information up to time N, the resulting forecasts are called *out-of-sample* forecasts. Economists call them *ex-ante* forecasts.

One difficulty for the analysts is that there is no immediate way of calibrating these forecasts except by waiting for future observations to become available. Thus it is sometimes helpful to check the forecasting ability of the model using data already at hand. This can be done in various ways. If the model is fitted to all the data and then used to 'forecast' the data already used in fitting the model, then the forecasts are sometimes called *in-sample forecasts*, even though they are not genuine forecasts. The one-step-ahead in-sample forecasts are in fact the *residuals* – see Section 3.5.3. An alternative procedure (see Section 8.5.4) is to split the data into two portions; the first, sometimes called the *training set*, is used to fit the model, while the second portion, sometimes called the *test set*, is used for calibration purposes to check forecasts made by the model. The properties of the resulting forecasts are more like those of real forecasts than the residuals. Of course, if one really believes one has the 'true' model, then the properties of residuals and forecast errors should be similar, but, we see in later chapters, that, in practice, out-of-sample forecasts are generally not as accurate as would be expected from in-sample fit. Fortunately, many of the comparative forecasting studies, which have been reported in the literature, do use a test set for making comparisons. Even so, there are many ways in which forecasts can be 'improved' by procedures which are of dubious validity and the reader is strongly advised to check that results on comparative forecast accuracy really do relate to forecasts made under similar conditions. In particular, if all forecasts are meant to be out-of-sample, then the different forecasting methods being compared should only use historical data in the information set.

There are several ways in which forecasts can be unfairly 'improved'. They include:

1. Fitting the model to all the data including the test set.
2. Fitting several different models to the training set, and then choosing the model which gives the best 'forecasts' of the test set. The selected model is then used (again) to produce forecasts of the test set, even though the latter has already been used in the modelling process.
3. Using the known test-set values of 'future' observations on the explanatory variables in multivariate forecasting. This will obviously improve forecasts of the dependent variable in the test set, but these future values will not of course be known at the time the forecast is supposedly made (though in practice the 'forecast' is made at a later

date). Economists call such forecasts *ex-post* forecasts to distinguish them from *ex-ante* forecasts. The latter, being genuinely out-of-sample, use forecasts of future values of explanatory variables, where necessary, to compute forecasts of the response variable – see also Section 5.1.2. *Ex-post* forecasts can be useful for assessing the effects of explanatory variables, provided the analyst does not pretend that they are genuine out-of-sample forecasts.

In my experience, it often seems to be the case that, when one method appears to give much better forecasts than alternatives, then the 'good' method has some unfair advantage. It is therefore unfortunate that some published empirical studies do not provide sufficient information to see exactly how forecasts were computed, and, in particular, to assess if they are genuinely out-of-sample. When this happens, the suspicion remains that the results are compromised. Of course, there will sometimes be good reasons for computing alternatives to genuine out-of-sample forecasts. For example, in scenario forecasting, the analyst wants to assess the effect of making different assumptions about future values of explanatory variables. This sort of exercise is perfectly reasonable provided that the assumptions are clearly stated and the forecaster realizes that it will not be fair to compare such results with forecasts based solely on past data.

1.5 Brief overview of relevant literature

This section gives a brief review of books on time-series analysis and forecasting. Some help with scanning research-level journals for articles (papers) on forecasting is also given. Additional references to more specialized books and papers are given throughout the book.

General introductory books on time-series analysis include Brockwell and Davis (1996), Chatfield (1996a), Diggle (1990), Harvey (1993), Janacek and Swift (1993), Kendall and Ord (1990) and Wei (1990). More advanced books include the comprehensive two-volume treatise by Priestley (1981) which is particularly strong on spectral analysis, multivariate time series and non-linear models. The fourth edition of Volume 3 of Kendall and Stuart (Kendall, Stuart and Ord, 1983) is also a valuable reference source, but earlier editions are now rather dated. Other intermediate to advanced books include Anderson (1971), Brillinger (1981), Brockwell and Davis (1991) and Fuller (1996). The books by Enders (1995), Hamilton (1994), Harvey (1990) and Mills (1990, 1999) are more suitable for the reader with an econometric background.

The famous book by Box and Jenkins (1970) describes an approach to time-series analysis, forecasting and control which is based on a class of linear stochastic processes, called ARIMA models. The revised edition published in 1976 was virtually unchanged, but the third edition (Box et al., 1994) with G. Reinsel as co-author is a substantial revision of earlier editions. In particular, Chapters 12 and 13 have been completely rewritten and include new material on topics such as intervention analysis, outlier detection and process control. We therefore generally refer to the new third edition. However, readers with an earlier edition will find that Chapters 1 to 11 of the new edition retain the spirit and a similar structure to the old one, albeit with some revisions such as new material on ARMA model estimation and on testing for unit roots. This book covers ARIMA models in Section 3.1 and gives a brief description of the Box-Jenkins approach in Section 4.2.2. Readers with little experience of ARIMA modelling, who want further details, may be advised to read Vandaele (1983), or one of the introductory time-series texts given above, rather than Box et al. (1994).

There are a number of books that are targeted more towards forecasting, rather than general time-series analysis. Granger and Newbold (1986) is a good general book on the topic, especially for applications in economics. Some other general texts on time-series forecasting include Abraham and Ledolter (1983) and Montgomery et al. (1990). The texts by Bowerman and O'Connell (1987), Diebold (1998), Franses (1998), and Makridakis et al. (1998) are aimed more at business and economics students. There is a useful collection of up-to-date review articles in Armstrong (2001). Some important, more specialized, books include Harvey's (1989) book on *structural models* and West and Harrison's (1997) book on *dynamic linear models*, which is written from a Bayesian viewpoint. Pole et al. (1994) give some case studies using the latter approach together with a software package called BATS.

The two main journals devoted to forecasting are the *International Journal of Forecasting* (sponsored by the International Institute of Forecasters and published by North-Holland) and the *Journal of Forecasting* (published by Wiley). Papers on forecasting can also appear in many other statistical, management science, econometric and operational research journals including the *Journal of Business and Economic Statistics*, *Management Science* and the *Journal of Econometrics*. A brief general review of recent developments in time-series forecasting is given by Chatfield (1997). Keeping up with the literature in an interdisciplinary subject like forecasting is difficult. Abstract journals may help and 'word of mouth' at specialist conferences is also invaluable.