2

Getting Started

Although exponential smoothing methods have been around since the 1950s, a modeling framework incorporating stochastic models, likelihood calculations, prediction intervals, and procedures for model selection was not developed until relatively recently, with the work of Ord et al. (1997) and Hyndman et al. (2002). In these (and other) papers, a class of state space models has been developed that underlies all of the exponential smoothing methods.

In this chapter, we provide an introduction to the ideas underlying exponential smoothing and the associated state space models. Many of the details will be skipped over in this chapter, but will be covered in later chapters.

Figure 2.1 shows the four time series from Fig. 1.1, along with point forecasts and 80% prediction intervals. These were all produced using exponential smoothing state space models. In each case, the particular models and all model parameters were chosen automatically with no intervention by the user. This demonstrates one very useful feature of state space models for exponential smoothing—they are easy to use in a completely automated way. In these cases, the models were able to handle data exhibiting a range of features, including very little trend, strong trend, no seasonality, a seasonal pattern that stays constant, and a seasonal pattern with increasing variation as the level of the series increases.

2.1 Time Series Decomposition

It is common in business and economics to think of a time series as a combination of various components such as the trend (T), cycle (C), seasonal (S), and irregular or error (E) components. These can be defined as follows:

Trend (T):	The long-term direction of the series
Seasonal (S):	A pattern that repeats with a known periodicity (e.g., 12 months per year, or 7 days per week)

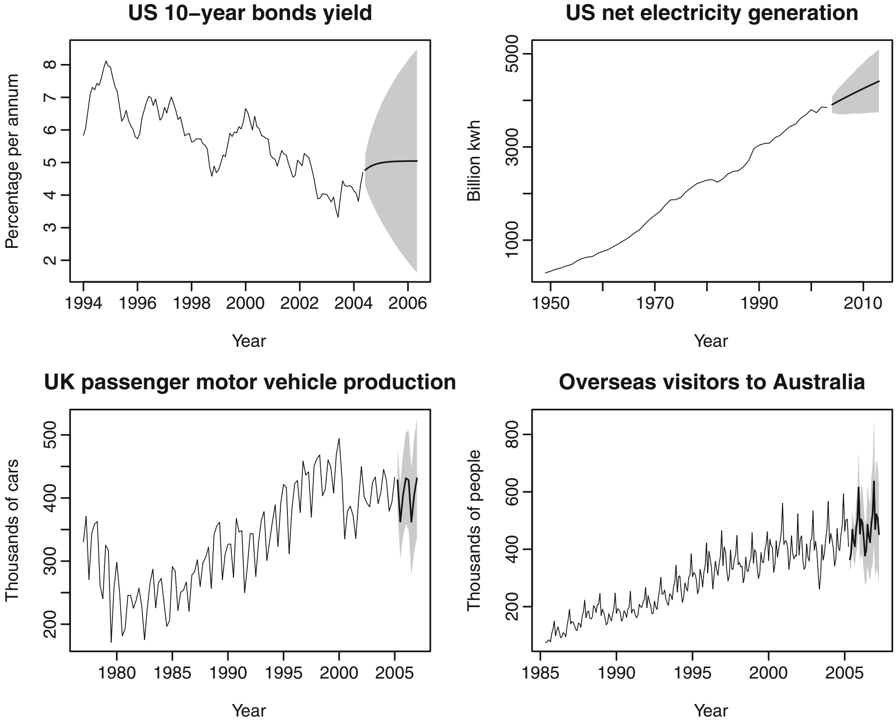

Fig. 2.1. Four time series showing point forecasts and 80% prediction intervals obtained using exponential smoothing state space models.

Cycle (C):	A pattern that repeats with some regularity but with unknown and changing periodicity (e.g., a business cycle)
Irregular or error (E):	The unpredictable component of the series

In this monograph, we focus primarily upon the three components T, S and E. Any cyclic element will be subsumed within the trend component unless indicated otherwise.

These three components can be combined in a number of different ways. A purely additive model can be expressed as

$$y = T + S + E,$$

where the three components are added together to form the observed series. A purely multiplicative model is written as

$$y = T \times S \times E,$$

where the data are formed as the product of the three components. A *seasonally adjusted* series is then formed by extracting the seasonal component from the data, leaving only the trend and error components. In the additive model, the seasonally adjusted series is $y - S$, while in the multiplicative model, the seasonally adjusted series is y/S. The reader should refer to Makridakis et al. (1998, Chap. 4) for a detailed discussion of seasonal adjustment and time series decomposition.

Other combinations, apart from simple addition and multiplication, are also possible. For example,

$$y = (T + S) \times E$$

treats the irregular component as multiplicative but the other components as additive.[1]

2.2 Classification of Exponential Smoothing Methods

In exponential smoothing, we always start with the trend component, which is itself a combination of a level term (ℓ) and a growth term (b). The level and growth can be combined in a number of ways, giving five future trend types. Let T_h denote the forecast trend over the next h time periods, and let ϕ denote a damping parameter ($0 < \phi < 1$). Then the five trend types or growth patterns are as follows:

$$\begin{aligned}
\text{None:}\quad & T_h = \ell \\
\text{Additive:}\quad & T_h = \ell + bh \\
\text{Additive damped:}\quad & T_h = \ell + (\phi + \phi^2 + \cdots + \phi^h)b \\
\text{Multiplicative:}\quad & T_h = \ell b^h \\
\text{Multiplicative damped:}\quad & T_h = \ell b^{(\phi+\phi^2+\cdots+\phi^h)}
\end{aligned}$$

A damped trend method is appropriate when there is a trend in the time series, but one believes that the growth rate at the end of the historical data is unlikely to continue more than a short time into the future. The equations for damped trend do what the name indicates: dampen the trend as the length of the forecast horizon increases. This often improves the forecast accuracy, particularly at long lead times.

Having chosen a trend component, we may introduce a seasonal component, either additively or multiplicatively. Finally, we include an error, either additively or multiplicatively. Historically, the nature of the error component has often been ignored, because the distinction between additive and multiplicative errors makes no difference to point forecasts.

If the error component is ignored, then we have the fifteen exponential smoothing methods given in the following table. This classification of methods originated with Pegels' (1969) taxonomy. This was later extended by

[1] See Hyndman (2004) for further discussion of the possible combinations of these components.

Trend component	Seasonal component		
	N (None)	A (Additive)	M (Multiplicative)
N (None)	N,N	N,A	N,M
A (Additive)	A,N	A,A	A,M
A_d (Additive damped)	A_d,N	A_d,A	A_d,M
M (Multiplicative)	M,N	M,A	M,M
M_d (Multiplicative damped)	M_d,N	M_d,A	M_d,M

Gardner (1985), modified by Hyndman et al. (2002), and extended again by Taylor (2003a), giving the fifteen methods in the above table.

Some of these methods are better known under other names. For example, cell (N,N) describes the simple exponential smoothing (or SES) method, cell (A,N) describes Holt's linear method, and cell (A_d,N) describes the damped trend method. Holt-Winters' additive method is given by cell (A,A), and Holt-Winters' multiplicative method is given by cell (A,M). The other cells correspond to less commonly used but analogous methods.

For each of the 15 methods in the above table, there are two possible state space models, one corresponding to a model with additive errors and the other to a model with multiplicative errors. If the same parameter values are used, these two models give equivalent point forecasts although different prediction intervals. Thus, there are 30 potential models described in this classification.

We are careful to distinguish exponential smoothing *methods* from the underlying state space *models*. An exponential smoothing method is an algorithm for producing point forecasts only. The underlying stochastic state space model gives the same point forecasts, but also provides a framework for computing prediction intervals and other properties. The models are described in Sect. 2.5, but first we introduce the much older point-forecasting equations.

2.3 Point Forecasts for the Best-Known Methods

In this section, a simple introduction is provided to some of the best-known exponential smoothing methods—simple exponential smoothing (N,N), Holt's linear method (A,N), the damped trend method (A_d,N) and Holt-Winters' seasonal method (A,A and A,M). We denote the observed time series by $y_1, y_2, \ldots, y_n$. A forecast of y_{t+h} based on all the data up to time t is denoted by $\hat{y}_{t+h|t}$. For one-step forecasts, we use the simpler notation $\hat{y}_{t+1} \equiv \hat{y}_{t+1|t}$. Usually, forecasts require some parameters to be estimated; but for the sake of simplicity it will be assumed for now that the values of all relevant parameters are known.