

# Business Forecasting Methods

Rob J Hyndman\*

## 1 Forecasting, planning and goals

Forecasting is a common statistical task in business where it helps inform decisions about scheduling of production, transportation and personnel, and provides a guide to long-term strategic planning. However, business forecasting is often done poorly and is frequently confused with planning and goals. They are three different things.

**Forecasting** is about predicting the future as accurately as possible, given all the information available including historical data and knowledge of any future events that might impact the forecasts.

**Goals** are what you would like to happen. Goals should be linked to forecasts and plans, but this does not always occur. Too often, goals are set without any plan for how to achieve them, and no forecasts for whether they are realistic.

**Planning** is a response to forecasts and goals. Planning involves determining the appropriate actions that are required to make your forecasts match your goals.

Forecasting should be an integral part of the decision-making activities of management, as it can play an important role in many areas of a company. Modern organizations require short-, medium-, and long-term forecasts, depending on the specific application.

**Short-term forecasts** are needed for scheduling of personnel, production and transportation. As part of the scheduling process, forecasts of demand are often also required.

**Medium-term forecasts** are needed to determine future resource requirements in order to purchase raw materials, hire personnel, or buy machinery and equipment.

**Long-term forecasts** are used in strategic planning. Such decisions must take account of market opportunities, environmental factors and internal resources.

An organization needs to develop a forecasting system involving several approaches to predicting uncertain events. Such forecasting systems require the development of expertise in identifying forecasting problems, applying a range of forecasting methods, selecting appropriate methods for each problem, and evaluating and refining forecasting methods over time. It is also important to have strong organizational support for the use of formal forecasting methods if they are to be used successfully.

## 2 Commonly used methods

Throughout the 20th century, businesses tended to use relatively simple forecasting methods that are often not based on statistical modelling. However, the use of statistical forecasting has been growing over the last 20 years, and more recently, some machine learning methods have become popular.

\* Naive \* SES, Holt, HW \* ETS \* ARIMA \* Linear regression \* Nonparametric additive models \* Dynamic regression \* Distributed lag models \* DeepAR and other NN-based methods \* XGboost?

---

\*Professor of Statistics, Department of Econometrics & Business Statistics, Monash University, Australia. For biography, see the entry **Forecasting overview**.

## Time series methods

Let the historical time series data be denoted by  $y_1, \dots, y_T$ , and the forecast of  $y_{T+h}$  be given by  $\hat{y}_{T+h|T}$ ,  $h > 0$ .

- Naïve forecasting is where the forecast of all future values of a time series are set to be equal to the last observed value:  $\hat{y}_{T+h|T} = y_T$ ,  $h = 1, 2, \dots$ . If the data follow a random walk process ( $y_t = y_{t-1} + e_t$  where  $e_t$  is white noise — a series of iid random variables with zero mean), then this is the optimal method of forecasting. Consequently, it is popular for stock price and stock index forecasting, and for other time series that measure the behaviour of a market that can be assumed to be efficient.
- Simple exponential smoothing was developed in the 1950s (Brown, 1959) and has been widely used ever since. Forecasts can be computed recursively as each new data point is observed:

$$\hat{y}_{t+1|t} = \alpha y_t + (1 - \alpha) \hat{y}_{t|t-1},$$

where  $0 < \alpha < 1$ . (Longer-term forecasts are constant:  $\hat{y}_{t+h|t} = \hat{y}_{t+1|t}$ ,  $h \geq 2$ .) Consequently, only the most recent data point and most recent forecast need to be stored. This was an attractive feature of the method when computer storage was expensive. The method has proved remarkably robust to a wide range of time series, and is optimal for several processes including the ARIMA(0,1,1) process (Chatfield et al., 2001).

- Holt's linear method (Holt, 1957) is an extension of simple exponential forecasting that allows a locally linear trend to be extrapolated. Forecasts are given by  $\hat{y}_{t+h|t} = \ell_t + hb_t$  where

$$\begin{aligned}\ell_t &= \alpha y_t + (1 - \alpha)(\ell_{t-1} + b_{t-1}), \\ b_t &= \beta(\ell_t - \ell_{t-1}) + (1 - \beta)b_{t-1},\end{aligned}$$

and the two parameters  $\alpha$  and  $\beta$  must lie in  $[0, 1]$ . Here  $\ell_t$  denotes the level of the series and  $b_t$  the slope of the trend at time  $t$ .

- For seasonal data, a popular method is the Holt-Winters' method, also introduced in Holt (1957), which extends Holt's method to include seasonal terms. Then  $\hat{y}_{t+h|t} = \ell_t + hb_t + s_{t-m+h_m^+}$ , where

$$\begin{aligned}\ell_t &= \alpha(y_t - s_{t-m}) + (1 - \alpha)(\ell_{t-1} + b_{t-1}), \\ b_t &= \beta(\ell_t - \ell_{t-1}) + (1 - \beta)b_{t-1}, \\ s_t &= \gamma(y_t - \ell_t) + (1 - \gamma)s_{t-m},\end{aligned}$$

$h_m^+ = [(h - 1) \bmod m] + 1$ , and the three parameters  $\alpha$ ,  $\beta$  and  $\gamma$  all lie in  $[0, 1]$ .

There is also a multiplicative version of the Holt-Winters' method, and damped trend versions of both Holt's linear method and Holt-Winters' method (Hyndman and Athanasopoulos, 2021). None of these methods are explicitly based on underlying time series models, and as a result the estimation of parameters and the computation of prediction intervals is often not done. However, all the above methods have been shown to be optimal for some state space models (Hyndman et al., 2008), and maximum likelihood estimation of parameters, statistical model selection and computation of prediction intervals is now becoming more widespread.

Other time series models sometimes used in business forecasting include ARIMA models, GARCH models (especially in finance), structural models and neural networks.

## Explanatory models for forecasting

The use of explanatory models in business forecasting does not have such a long history as the use of time series methods.

- Linear regression modelling is now widely used (e.g., Pardoe, 2006) where a variable to be forecast is modelled as linear combination of potential input variables:

$$y_t = \sum_{j=1}^J c_j x_{j,t} + e_t,$$

where  $e_t$  denotes an iid error term with zero mean. An interesting application of regression model to forecasting is given by Byron and Ashenfelter (1995) who use a simple regression model to predict the quality of a Grange wine using simple weather variables. However, it is far more common for regression modelling to be used to explain historical variation than for it to be used for forecasting purposes.

- In some domains, the use of nonparametric additive models for forecasting is growing (e.g., Hyndman and Fan, 2010). Here, the model is often of the form

$$y_t = \sum_{j=1}^J f_j(x_{j,t}) + e_t,$$

where  $f_j$  is a smooth nonlinear function to be estimated nonparametrically.

- In advertising, there is a well-developed culture of using distributed lag regression models (e.g., Hanssens et al., 2001) such as

$$y_t = \sum_{j=1}^J \alpha \lambda^j x_{t-j} + e_t,$$

where  $x_t$  denotes advertising expenditure in month  $t$ ,  $0 < \lambda < 1$  and  $\alpha > 0$ .

## Data mining methods for business forecasting

Outside of traditional statistical modelling, an enormous amount of forecasting is done using data mining methods. Most of these methods have no formal statistical model, prediction intervals are not computed, and there is limited model checking. But some of the data-mining methods have proven powerful predictors in some contexts, especially when there is a vast quantity of available data. Predictive methods include neural networks, support vector machines and regression trees. Many of the most best-known business predictive algorithms are based on data-mining methods including the prediction of Netflix ratings (see **Data Mining Time Series Data, Forecasting Principles** ).

## References

- Brown, R. G. (1959), *Statistical forecasting for inventory control*, McGraw-Hill, New York.
- Byron, R. P. and Ashenfelter, O. (1995), 'Predicting the quality of an unborn Grange', *The Economic Record* **71**(212), 40–53.
- Chatfield, C., Koehler, A. B., Keith Ord, J. and Snyder, R. D. (2001), 'A new look at models for exponential smoothing', *Journal of the Royal Statistical Society: Series D* **50**(2), 147–159.
- Hanssens, D. M., Parsons, L. J. and Schultz, R. L. (2001), *Market Response Models: Econometric and Time Series Analysis*, 2nd edn, Kluwer Academic Publishers, Boston, MA.
- Holt, C. C. (1957), Forecasting trends and seasonals by exponentially weighted averages, O.N.R. Memorandum 52/1957, Carnegie Institute of Technology.
- Hyndman, R. J. and Athanasopoulos, G. (2021), *Forecasting: principles and practice*, 3rd edition edn, OTexts, Melbourne, Australia.  
**URL:** [OTexts.org/fpp3](https://otexts.org/fpp3)

- Hyndman, R. J. and Fan, S. (2010), 'Density forecasting for long-term peak electricity demand', *IEEE Transactions on Power Systems* **25**(2), 1142–1153.
- Hyndman, R. J., Koehler, A. B., Ord, J. K. and Snyder, R. D. (2008), *Forecasting with exponential smoothing: the state space approach*, Springer-Verlag, Berlin.  
**URL:** [www.exponentialsMOOTHING.net](http://www.exponentialsMOOTHING.net)
- Pardoe, I. (2006), *Applied regression modeling: a business approach*, John Wiley & Sons, Hoboken, NY.