



ELSEVIER

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

International Journal of Forecasting 22 (2006) 403–410

international journal
of forecasting

www.elsevier.com/locate/ijforecast

Book reviews

Ian T. Jolliffe, David B. Stephenson, *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, Chichester, John Wiley and Sons, 2003, ISBN: 0-471-49759-2

One could easily assume from the title that the book *Forecast verification: A practitioner's guide in atmospheric science* was primarily written for those interested in meteorological research. This could not be further from the truth. While many of the forecast evaluation techniques critiqued in the book have their genesis in the meteorological literature, the clear development and exposition of the evaluation of both deterministic and probabilistic forecasts is a feature of the book. Along with excellent referencing of a wealth of reported research, the end result makes it a mandatory acquisition for those who have an interest in forecast evaluation that has been spawned from more diverse backgrounds.

The nine chapters of the book, all written by different contributors, have been logically sequenced and carefully edited. Remarkably, given the sophistication and interrelatedness of much of the content, there is clarity and consistency in the exposition of ideas from one chapter to the next. The consistency of definitions and technical jargon associated with a range of evaluation techniques introduced early in the book, and further developed in later chapters, combined with constant cross-referencing between chapters, gives a uniformity of presentation. The reader is left with the impression that the writing is that of a single author. Tables are succinct and figures used for illustrative purposes uncomplicated.

Chapter 1 introduces the reader to the history and benefits of forecast verification. Importantly, it deals with how to measure the relative quality of a forecast by establishing a suitable baseline to score forecasting skill and its significance from the perspective of adding value. The importance of avoiding biases in

skill scores, along with data quality considerations, make up the remainder of the introduction. Chapter 2 covers the various statistical concepts and definitions underpinning forecast verification which are used in subsequent chapters. While this chapter contains material that would be considered basic by most experienced forecasters, the treatment of the Murphy–Winkler framework for forecast verification based on the joint distribution of forecasts and observations is a section that provides an important platform for later chapters.

Chapter 3 is concerned with the reliable measurement of the quality of binary events. After introducing and analysing a number of available performance measures of forecasts of binary events, there follows an excellent treatment of Signal Detection Theory (SDT). The advantages offered by SDT for the evaluation of probabilistic forecasts of an event by choosing a range of probability decision thresholds and then considering the set of deterministic binary forecasts that correspond to those thresholds are highlighted. The distinction between the discrimination stage of a forecast system that assesses the degree to which current evidence favours one alternative over another, and the decision processes that lead to the acceptance of one alternative over another based on the strength of this evidence, is facilitated by the use of an analysis tool known as the receiver (relative) operating characteristic (ROC) curve. ROC curve analysis is an extremely useful approach for evaluation of probabilistic forecasts. The development of this technique in this chapter, and later in Chapter 7 where it is extended, is guaranteed to empower potential users.

Chapter 4 is predominantly concerned with the desirable attributes and use of skill scores for ordinal categorical event verification of forecast problems with more than two categories. While the focus is on discrete events, the multi-categorical

use of skill scores is also defined in terms of ranges of continuous variables. On the other hand, Chapter 5 presents methods for the verification of real continuous scalar quantities that are, increasingly, the output from contemporary forecasting systems. Verification criteria based on first-order moments that include mean systematic error or bias and mean absolute error, along with criteria based on second order moments such as mean squared error and correlation, are discussed in detail. A discussion of skill issues associated with bias correction of these criteria makes a further important contribution.

Forecast evaluation of spatial fields that involve the same parameter over a range of geographic locations, rather than some averaging process over time, is covered in Chapter 6. Spatial fields are revealed by partitioning the data into subsets of forecasts at each location for a given time. Spatial averaging gives a score for a particular time, calculated over all spatial points and to which most of the standard verification measures can be applied. Summation is over spatial points rather than time. After dealing with the measures commonly used in the spatial domain, the assessment of models is addressed through principal component analysis and signal detection analysis.

Chapter 7 deals with probabilistic and ensemble forecasts. While probabilistic forecasts can be generated by forecasters using a range of methods that consider a wide range of forecast information, an ensemble forecast is an estimate of the forecast probability of an event based on the set of deterministic forecasts that are valid at the same time. It is the fraction of forecasts predicting the event among all the forecasts considered at any time, while assuming that they are all independent realizations of the same underlying random process. This chapter extends the ideas introduced in Chapter 3. After defining and further discussing the main attributes of probabilistic forecasts, namely, reliability and resolution, verification techniques for binary events like the reliability curve and the Brier Score are assessed. Verification based on decision probability thresholds and the ROC curve is established as a measure of resolution that provides similar qualitative information to the Brier score. Further, verification statistics for measuring the performance of probabilistic forecasts with respect to the attributes of reliability and resolution

are generalized to multi-outcome events. For an ensemble forecast system, reliability is a function of the statistical consistency of the forecasts from the system with observations of it. As a result, most of the verification tools applied to ensemble forecasts are designed to evaluate the statistical consistency of the forecasts making up the ensemble. By exploring and summarizing their basic statistical properties, weak points of the ensemble system can be exposed. Following a detailed discussion of these measures, a section dealing with some of the limitations of probabilistic and ensemble verification concludes this chapter.

Chapter 8 concentrates on economic value and its relationship to the quality measures of forecast skill. Given that a forecaster has available a number of alternative courses of action, the task of the decision maker should be to choose the appropriate action that minimises the expected loss or maximises the expected benefit. This chapter introduces the reader to the implementation of economic decision models using both deterministic and probabilistic forecasts by highlighting the link between user value and verification measures via the (static) cost–loss economic decision model.

Chapter 9 contains a review of the key concepts of forecast verification and gives an overview of its use in other disciplines, including statistics, finance and economics, environmental and earth sciences, as well as medical and clinical studies. Key seminal papers and web sites are suggested in order to direct the reader to a further elaboration of verification techniques within these disciplines. The chapter concludes with a short discussion of future directions that can and may be taken in forecast verification, along with an appeal to forecasters for clear documentation and public availability of previous and on-going forecast performance. This, it is hoped, will enhance future forecasting improvements through examination and feedback from third parties.

This book will prove extremely useful for practitioners and academics alike. While it certainly makes a rigorous contribution to the understanding and implementation of verification techniques for categorical and probabilistic forecasts, in the main it should also satisfy the demand for these techniques made by graduate students.

Maxwell Stevenson
School of Business, The University of Sydney,
Sydney N.S.W. 2006 Australia
E-mail address: m.stevenson@econ.usyd.edu.au.

doi:10.1016/j.ijforecast.2005.11.002

Hans Levenbach, James P. Cleary, *Forecasting: Practice and process for demand management*, Duxbury, Belmont: Thomson, 2006, 622 pp., ISBN: 0-534-26268-6, Hardcover (software enclosed)

This is a very comprehensive account covering many important aspects of demand forecasting and management. The structure of the book is very logical indeed, commencing with a clear introductory section on the objectives of this work, target audience and appropriate use of the material presented. The book is aimed at practicing forecasters and students even though it is clear that managers and end-users of forecasts will also find in this book a comprehensive treatment of how to evaluate basic demand forecasting approaches. Regarding this latter group of potential readers, there are some very useful suggestions for running courses and seminars for the purpose of enhancing the technical skills of practitioners performing demand forecasting functions in corporate environments. Four courses are recommended with a clear indication of the relevant chapters/sections that should be covered: structured approach to demand forecasting, demand forecasting and market analysis, time series and smoothing techniques, data analysis and modelling demand.

The principal unifying theme of this work is the presentation of demand forecasting as a process rather than a series of disconnected techniques. Demand forecasting, within the context of this book, means that the firm predicts the *right* amount of the *right* product to be in the *right* place at the *right* time for the *right* price, which is one of the underpinnings of what is now known as demand forecasting and replenishment planning for the supply chain. In those respects the book is to be distinguished from other typical textbooks that cover business forecasting in general.

The authors take the view that the focus is nowadays directed more to forecasting the disaggregated elements of product demand for supplying

warehouses, distributors, channels and consumers than to aggregates drive by economic and financial factors. Consequently, the most widely accepted relevant quantitative methods are presented. The development of this book progresses from the basic, most widely used techniques to the more sophisticated, less practiced methods and hypothesis testing theory is intentionally de-emphasised in favour of confidence interval estimation.

Qualitative/subjective methods are not covered in this book. A brief reference to panel consensus, the Delphi method, historical analogues, etc., is made in chapter 2 directing the reader for more information to other textbooks. DeLurgio (1998) and Hanke and Reitsch (1998) are used as extra points of reference for business forecasting in the majority of chapters. Moreover, some of the “esoteric” methods, such as neural networks, vector auto-regression and GARCH (Generalised Auto-Regressive Conditionally Heteroscedastic) models are also not included since they appear to be more relevant to applications to finance than demand planning.

Many practical examples are given in the book and data sets from a variety of sources have been used throughout to make certain points or illustrate a particular technique. Very good supplementary material is also included in the CD enclosed. According to the authors the failure of many forecasting efforts begins with flaws in the quality and handling of data rather than in the lack of modelling sophistication. Thus the book places greater emphasis on data analytic methodology (much of it intuitive and graphical) as a key to improved forecasting. The book deals with exploratory data analysis along with confirmatory modelling and consequently it emphasises techniques for which a “reasonable” amount of data is available or can be collected. As such, technological forecasting is not treated in this book. Nevertheless, new product forecasting, for which adequate data are rarely available, is covered.

The book is divided into six parts. Part 1 (chapters 1 and 2) introduces the forecasting process along with a broad classification of forecasting techniques.

Part 2 (chapters 3–5) deals with basic statistical concepts for exploring time series characteristics with a very comprehensive chapter 4 on the characteristics of time series. In the introduction to part 2 the authors caution the reader that “...most testing procedures