ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

de Gooijer, Jan G.; Hyndman, Rob J.

Working Paper

25 Years of IIF Time Series Forecasting: A Selective Review

Tinbergen Institute Discussion Paper, No. 05-068/4

Provided in Cooperation with:
Tinbergen Institute, Amsterdam and Rotterdam

Suggested Citation: de Gooijer, Jan G.; Hyndman, Rob J. (2005) : 25 Years of IIF Time Series Forecasting: A Selective Review, Tinbergen Institute Discussion Paper, No. 05-068/4, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:
https://hdl.handle.net/10419/86645

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

TI 2005-068/4
Tinbergen Institute Discussion Paper

25 Years of IIF Time Series Forecasting: A Selective Review

Jan G. De Gooijer[1]
Rob J. Hyndman[2]

[1] *University of Amsterdam, and Tinbergen Institute,*
[2] *Monash University, Australia.*

Tinbergen Institute
The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Please send questions and/or remarks of non-scientific nature to driessen@tinbergen.nl.
Most TI discussion papers can be downloaded at http://www.tinbergen.nl.

25 Years of IIF Time Series Forecasting: A Selective Review[1]

Jan G De Gooijer
Department of Quantitative Economics
University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands
Telephone: +31–20–525–4244; Fax: +31–20–525–4349
Email: j.g.degooijer@uva.nl

Rob J Hyndman
Department of Econometrics and Business Statistics,
Monash University, VIC 3800, Australia.
Telephone: +61–3–9905–2358; Fax: +61–3–9905–5474
Email: Rob.Hyndman@buseco.monash.edu

[1]Also available as Working paper 12/05 Department of Econometrics & Business Statistics, Monash University, Australia; and at REPEC; htttp://ideas.repec.org/p/msh/ebswps/2005-12.html

25 Years of IIF Time Series Forecasting: A Selective Review

Abstract: We review the past 25 years of time series research that has been published in journals managed by the International Institute of Forecasters (*Journal of Forecasting* 1982–1985; *International Journal of Forecasting* 1985–2005). During this period, over one third of all papers published in these journals concerned time series forecasting. We also review highly influential works on time series forecasting that have been published elsewhere during this period. Enormous progress has been made in many areas, but we find that there are a large number of topics in need of further development. We conclude with comments on possible future research directions in this field.

Keywords: Accuracy measures; ARCH model; ARIMA model; Combining; Count data; Densities; Exponential smoothing; Kalman filter; Long memory; Multivariate; Neural nets; Nonlinearity; Prediction intervals; Regime switching models; Robustness; Seasonality; State space; Structural models; Transfer function; Univariate; VAR.

1 Introduction

Twenty five years of the International Institute of Forecasters (IIF) is synonymous to 24 years (1982–2005) of scientific publications in two leading journals in the field. In 1982 the IIF set up the *Journal of Forecasting* (*JoF*), published with John Wiley & Sons. After a break with Wiley in 1985[2] the IIF decided to start with the *International Journal of Forecasting* (*IJF*), published with Elsevier since 1985. This paper provides a selective guide to the literature on time series forecasting, covering the period 1982–2005 and summarizing about 340 papers published under the "IIF-flag" out of a total of over 940 papers. The proportion of these time series forecasting papers as a function of the year of publication is shown in Figure 1, and we note that the behaviour of the series is rather stable.[3] The review also includes works (key papers and books) that have been highly influential to various developments in the field. The works referenced comprise 312 journal papers, and 13 books and monographs.

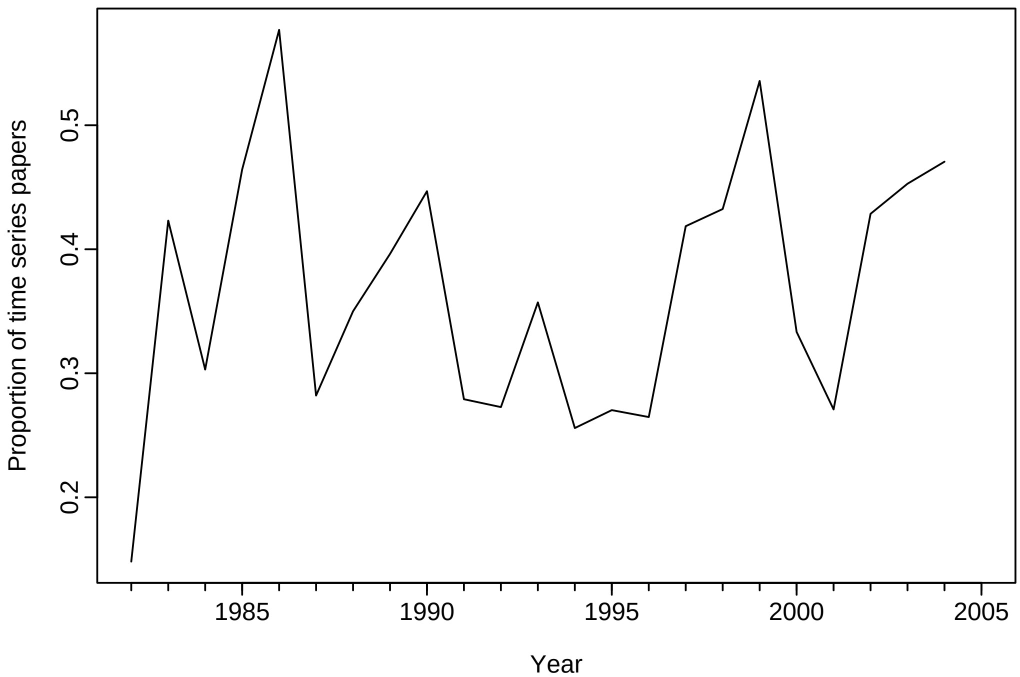

Figure 1: *Annual proportion of papers published by the IIF that concern time series forecasting.*

It was felt convenient to first classify the papers according to the models (e.g., exponential smoothing, ARIMA) introduced in the time series literature, rather than putting papers under a heading associated with a particular method. For instance, Bayesian methods in general can be applied to all models. Papers not concerning a particular model were then classified according to the various problems (e.g., accuracy measures, combining) they address. In only a few cases was a subjective decision on our part needed to classify a paper under a particular section heading. To facilitate a quick overview in a particular field, the papers are referenced

[2] The IIF was involved with *JoF* issue 14:1 (1985)

[3] *IJF* issue 7:4 (1991) was erroneously dated 1992. The time series forecasting papers published in this issue are assumed to be published in the year 1991.

in alphabetical order according to the section headings.

Determining what to include and what not to include in the list of references has been a problem. There may be papers that we have missed, and papers that are also referenced by other authors in this Silver Anniversary issue. For example, the ARCH/GARCH papers (Section 8) will probably be covered in more detail in the finance survey paper, and the vector error correction/cointegration papers discussed in Subsection 3.4 will also be covered in the econometrics survey paper. If an article did not belong to the mainstream of time series forecasting, it was included only if it was deemed important for forecasters to know about.

The review is not intended to be critical, but rather a (brief) historical and personal tour of the main developments. Still, a cautious reader may detect certain areas where the fruits of 25 years of intensive research interest has been limited. Conversely, clear explanations for many previously anomalous time series forecasting results have been provided by the end of 2005. Section 13 discusses some current research directions that hold promise for the future, but of course the list is far from exhaustive.

2 Exponential smoothing

2.1 Preamble

Twenty five years ago, exponential smoothing methods were considered a collection of ad hoc techniques for extrapolating various types of univariate time series. Although exponential smoothing methods were widely used in business and industry, they had received little attention from statisticians and did not have a well-developed statistical foundation. These methods originated in the 1950s and 1960s with the work of Brown (1959, 1963), Holt (1957, 2004) and Winters (1960). Pegels (1969) provided a simple but useful classification of the trend and the seasonal patterns depending on whether they are additive (linear) or multiplicative (nonlinear).

Exponential smoothing methods received a boost by two papers published in 1985, which laid the foundation for much of the subsequent work in this area. First, Gardner (1985) provided a thorough review and synthesis of work in exponential smoothing to that date, and extended Pegels' classification to include damped trend. This paper brought together a lot of existing work which stimulated the use of these methods and prompted a substantial amount of additional research. Later in the same year, Snyder (1985) showed that simple exponential smoothing (SES) could be considered as arising from an innovation state space model (i.e., a model with a single source of error). Although this insight went largely unnoticed at the time, in recent years it has provided the basis for a large amount of work on state space models underlying exponential smoothing methods.

Most of the work since 1985 has involved studying the empirical properties of the methods (e.g., Bartolomei & Sweet, 1989; Makridakis & Hibon, 1991), proposals for new methods of estimation or initialization (Ledolter & Abraham, 1984), evaluation of the forecasts (Sweet & Wilson, 1988; McClain, 1988), or has concerned statistical models that can be considered to underly the methods (e.g., McKenzie, 1984). The damped multiplicative methods of Taylor (2003) provide the only genuinely new exponential smoothing methods over this period. There have, of course, been numerous studies applying exponential smoothing methods in various contexts including computer components (Gardner, 1993), air passengers (Grubb & Masa, 2001) and production planning (Miller & Liberatore, 1993).

Hyndman et al.'s (2002) taxonomy (extended by Taylor, 2003) provides a helpful notation in describing the various methods. Each method is denoted by one or two letters for the trend (N=none, A=additive, DA=damped additive, M=multiplicative and DM=damped multiplicative) and one for the seasonality (N=none, A=additive and M=multiplicative). Thus, there are 15 different methods, the best known of which are simple exponential smoothing (N-N), Holt's linear method (A-N), Holt-Winters' additive method (A-A) and Holt-Winters' multiplicative method (A-M).

2.2 Variations

Numerous variations on the original methods have been proposed. For example, Carreno & Madinaveitia (1990) and Williams & Miller (1999) proposed modifications to deal with discontinuities, and Rosas & Guerrero (1994) looked at exponential smoothing forecasts subject to one or more constraints. There are also variations in how and when seasonal components should be normalized. Lawton (1998) argued for renormalization of the seasonal indices at each time period, as it removes bias in estimates of level and seasonal components. Slightly different normalization schemes were given by Roberts (1982) and McKenzie (1986). Archibald & Koehler (2003) developed new renormalization equations that are simpler to use and give the same point forecasts as the original methods.

One useful variation, part way between SES and Holt's method, is SES with drift. This is equivalent to Holt's method with the trend parameter set to zero. Hyndman & Billah (2003) showed that this method was also equivalent to Assimakopoulos & Nikolopoulos's (2000) "Theta method" when the drift parameter is set to half the slope of a linear trend fitted to the data. The Theta method performed extremely well in the M3-competition, although why this particular choice of model and parameters is good has not yet been determined.

There has been remarkably little work in developing multivariate versions of the exponential smoothing methods. One notable exception is Pfeffermann & Allon (1989) who looked at Israeli tourism data.

2.3 State space models

Ord et al. (1997) built on the work of Snyder (1985) by proposing a class of innovation state space models which can be considered as underlying some of the exponential smoothing methods. Hyndman et al. (2002) and Taylor (2003) extended this to include all of the 15 exponential smoothing methods. In fact, Hyndman et al. (2002) proposed two state space models for each method, corresponding to the additive error and the multiplicative error cases. These models are not unique, and other related state space models for exponential smoothing methods are presented in Koehler et al. (2001) and Chatfield et al. (2001). It has long been known that some ARIMA models give equivalent forecasts to the linear exponential smoothing methods. The significance of the recent work on innovation state space models is that the nonlinear exponential smoothing methods can also be derived from statistical models.

2.4 Method selection

Gardner & McKenzie (1988) provided some simple rules based on the variances of differenced time series for choosing an appropriate exponential smoothing method. Tashman & Kruk (1996) compared these rules with others proposed by Collopy & Armstrong (1992) and an approach based on the BIC. Hyndman et al. (2002) also proposed an information criterion approach, but using the underlying state space models.

2.5 Robustness

The remarkably good forecasting performance of exponential smoothing methods has been addressed by several authors. Satchell & Timmermann (1995) and Chatfield et al. (2001) showed that SES is optimal for a wide range of data generating processes. In a simulation study, Hyndman (2001) showed that exponential smoothing performs better than ARIMA models because it is not so subject to model selection problems, particularly when data are non-normal.

2.6 Prediction intervals

One of the criticisms of exponential smoothing methods 25 years ago was that there was no way to produce prediction intervals for the forecasts. The first analytical approach to this problem was to assume the series were generated by deterministic functions of time plus white noise (Brown, 1963; Sweet, 1985; McKenzie, 1986; Gardner, 1985). If this was so, a regression model should be used rather than exponential smoothing methods; thus, Newbold & Bos (1989) strongly criticized all approaches based on this assumption.

Other authors sought to obtain prediction intervals via the equivalence between exponential smoothing methods and statistical models. Johnston & Harrison (1986) found forecast variances for the simple and Holt exponential smoothing methods for state space models with multiple sources of errors. Yar & Chatfield (1990) obtained prediction intervals for the additive Holt-Winters' method, by deriving the underlying equivalent ARIMA model. Approximate prediction intervals for the multiplicative Holt-Winters' method were discussed by Chatfield & Yar (1991) making the assumption that the one-step-ahead forecast errors are independent. Koehler et al. (2001) also derived an approximate formula for the forecast variance for the multiplicative Holt-Winters' method, differing from Chatfield & Yar (1991) only in how the standard deviation of the one-step-ahead forecast error is estimated.

Ord et al. (1997) and Hyndman et al. (2002) used the underlying innovation state space model to simulate future sample paths and thereby obtained prediction intervals for all the exponential smoothing methods. Hyndman et al. (2005) used state space models to derive analytical prediction intervals for 15 of the 30 methods, including all the commonly-used methods. They provide the most comprehensive algebraic approach to date for handling the prediction distribution problem for the majority of exponential smoothing methods.

2.7 Parameter space and model properties

It is common practice to restrict the smoothing parameters to the range 0 and 1. However, now that underlying statistical models are available, the natural (invertible) parameter space for the models can be used instead. Archibald (1990) showed that it is possible for smoothing parameters within the usual intervals to produce non-invertible models. Consequently, when forecasting, the impact of change in the past values of the series is non-negligible. Intuitively, such parameters produce poor forecasts and the forecast performance deteriorates. Lawton (1998) also discussed this problem.

3 ARIMA models

3.1 Preamble

In a series of articles, and a subsequent book, Box & Jenkins (1970, 1976) set out and illustrated a methodology for building autoregressive integrated moving average (ARIMA) forecasting models; see Newbold (1983) for an early survey. Their work has had an enormous impact on the theory and practice of modern time series analysis and forecasting.

Forecasting time series through univariate ARIMA models, transfer function (dynamic regression) models and multivariate (vector) ARIMA models has generated quite a few *IJF* papers. Often these studies were of an empirical nature, using one or more benchmark methods/models as a comparison. Without pretending to be complete, Table 1 gives a list of these studies. Naturally, some of these studies are more successful than others. In all cases, the forecasting experiences reported are valuable. They have also been the key to new developments which may be summarized as follows.

3.2 Univariate

If a time series is known to follow a univariate ARIMA model, forecasts using disaggregated observations are, in terms of MSE, at least as good as forecasts using aggregated observations. However, in practical applications there are other factors to be considered, such as missing values in disaggregated series. Both Ledolter (1989) and Hotta (1993) analysed the effect of the additive outlier on the forecast intervals when the ARIMA model parameters are estimated.

The problem of incorporating external (prior) information in the univariate ARIMA forecasts have been considered by Cholette (1982), Guerrero (1991) and de Alba (1993).

There are a number of methods (cf. Box & Jenkins, 1970, 1976) for estimating parameters of an ARMA model. Although these methods are equivalent asymptotically, in the sense that esti-

Data set	Forecast horizon	Benchmark	Reference
Univariate ARIMA			
Electricity load (minutes)	1–30 minutes	Wiener filter	Di Caprio et al. (1983)
Quarterly automobile insurance paid claim costs	8 quarters	log-linear regression	Cummins & Griepentrog (1985)
Daily federal funds rate	1 day	random walk	Hein & Spudeck (1988)
Quarterly macroeconomic data	1–8 quarters	Wharton model	Dhrymes & Peristiani (1988)
Monthly department store sales	1 month	simple exponential smoothing	Geurts & Kelly (1986, 1990); Pack (1990)
Monthly demand for telephone services	3 years	univariate state space	Grambsch & Stahel (1990)
Yearly population totals	20–30 years	demographic models	Pflaumer (1992)
Monthly tourism demand	1–24 months	univariate state space; multivariate state space	du Preez & Witt (2003)
Dynamic regression/Transfer function			
Monthly telecommunications traffic	1 month	univariate ARIMA	Layton et al. (1986)
Weekly sales data	2 years	n.a.	Leone (1987)
Daily call volumes	1 week	Holt-Winters	Bianchi et al. (1998)
Monthly employment levels	1–12 months	univariate ARIMA	Weller (1989)
Monthly and quarterly consumption of natural gas	1 month/ 1 quarter	univariate ARIMA	Liu & Lin (1991)
Monthly electricity consumption	1–3 years	univariate ARIMA	Harris & Liu (1993)
VARIMA			
Yearly municipal budget data	yearly (in-sample)	univariate ARIMA	Downs & Rocke (1983)
Monthly accounting data	1 month	regression, univariate, ARIMA, transfer function	Hillmer et al. (1983)
Quarterly macroeconomic data	1–10 quarters	judgmental methods, univariate ARIMA	Öller (1985)
Monthly truck sales	1–13 months	univariate ARIMA, Holt-Winters	Heuts & Bronckers (1988)
Monthly hospital patient movements	2 years	univariate ARIMA, Holt-Winters	Lin (1989)
Quarterly unemployment rate	1–8 quarters	transfer function	Edlund & Karlsson (1993)

Table 1: *A list of examples of real applications.*

mates tend to the same normal distribution, there are large differences in finite sample properties. In a comparative study of software packages, Newbold et al. (1994) showed that this difference can be quite substantial and, as a consequence, may influence forecasts. They recommended the use of full maximum likelihood. Kim (2003) considered parameter estimation and forecasting of AR models in small samples. He found that bias-corrected parameter estimators produce more accurate forecasts than the least squares estimator. Landsman & Damodaran (1989) presented evidence that the James-Stein ARIMA parameter estimator improves forecast accuracy relative to other methods, under a MSE loss criterion.

As an alternative to the univariate ARIMA methodology, Parzen (1982) proposed the ARARMA methodology. The key idea is that a time series is transformed from a long memory AR filter to a short-memory filter, thus avoiding the "harsher" differencing operator. In addition, a different approach to the 'conventional' Box-Jenkins identification step is used. In the M-competition (Makridakis et al., 1982), the ARARMA models achieved the lowest MAPE for longer forecast horizons. Meade (2000) compared the forecasting performance of an automated and non-automated ARARMA method.

Automatic univariate ARIMA modelling has been shown to produce one-step-ahead forecasting as accurate as those produced by competent modellers (Hill & Fildes, 1984; Libert, 1984; Poulos et al., 1987; Texter & Ord, 1989). Several software vendors have implemented automated time series forecasting methods (including multivariate methods); see, e.g., Geriner & Ord (1991), Tashman & Leach (1991) and Tashman (2000). Often these methods act as black boxes. The technology of expert systems (Mélard & Pasteels, 2000) can be used to avoid this problem.

Rather than adopting a single AR model for all forecast horizons Kang (2003) empirically investigated the case of using a multi-step ahead forecasting AR model selected separately for each horizon. The forecasting performance of the multi-step ahead procedure appears to depend on, among other things, optimal order selection criteria, forecast periods, forecast horizons, and the time series to be forecasted.

3.3 Transfer function

The identification of transfer function models can be difficult when there is more than one input variable. Edlund (1984) presented a two-step method for identification of the impulse response function when a number of different input variables are correlated. Koreisha (1983) established various relationships between transfer functions, causal implications and econometric model specification. Gupta (1987) identified the major pitfalls in causality testing. Using principal component analysis, a parsimonious representation of a transfer function model was suggested by del Moral & Valderrama (1997). Krishnamurthi et al. (1989) showed how more accurate estimates of the impact of interventions in transfer function models can be obtained by using a control variable.

3.4 Multivariate

The vector ARIMA (VARIMA) model is a multivariate generalization of the univariate ARIMA model. The population characteristics of VARMA processes appear to have been first derived by Quenouille (1957, 1968), although they were not routinely applied until the early 1980s. Since VARIMA models can accommodate assumptions on exogeneity and on contemporaneous relationships, they offered new challenges to forecasters and policy makers. Riise & Tjøstheim (1984) addressed the effect of parameter estimation on VARMA forecasts. Cholette & Lamy (1986) showed how smoothing filters can be built into VARMA models. The smoothing prevents irregular fluctuations in explanatory time series from migrating to the forecasts of the dependent series. To determine the maximum forecast horizon of VARMA processes, De Gooijer & Klein (1991) established the theoretical properties of cumulated multi-step-ahead forecasts and cumulated multi-step-ahead forecast errors. Lütkepohl (1986) studied the effects of temporal aggregation and systematic sampling on forecasting, assuming that the disaggregated (stationary) variable follows a VARMA process with unknown order. Later, Bidarkota (1998) considered the same problem but with the observed variables integrated rather than stationary.

Vector autoregressions (VARs) constitute a special case of the more general class of VARMA models. In essence, a VAR model is a fairly unrestricted (flexible) approximation to the reduced form of a wide variety of dynamic econometric models. VAR models can be specified in a number of ways. Funke (1990) presented five different VAR specifications and compared their forecasting performance using monthly industrial production series. Dhrymes & Thomakos (1998) discussed issues regarding the identification of structural VARs. Hafer & Sheehan (1989) showed the effect on VAR forecasts of changes in the model structure. Explicit expressions for VAR forecasts in levels are provided by Ariño & Franses (2000); see also Wieringa & Horváth (2005). Hansson et al. (2005) used a dynamic factor model as a starting point to obtain forecasts from parsimoniously parametrised VARs.

In general, VAR models tend to suffer from 'overfitting' with too many free insignificant parameters. As a result, these models can provide poor out-of-sample forecasts, even though within-sample fitting is good; see, e.g., Liu et al. (1994) and Simkins (1995). Instead of restricting some of the parameters in the usual way, Litterman (1986) and others imposed a prior distribution on the parameters expressing the belief that many economic variables behave like a random walk. BVAR models have been chiefly used for macroeconomic forecasting (Ashley, 1988; Kunst & Neusser, 1986; Artis & Zhang, 1990; Holden & Broomhead, 1990), for forecasting market shares (Ribeiro Ramos, 2003), for labor market forecasting (LeSage & Magura, 1991), for business forecasting (Spencer, 1993), or for local economic forecasting (LeSage, 1989). Kling & Bessler (1985) compared out-of-sample forecasts of several then-known multivariate time series methods, including Litterman's BVAR model.

The Engle-Granger (1987) concept of cointegration has raised various interesting questions regarding the forecasting ability of error correction models (ECMs) over unrestricted VARs and BVARs. Shoesmith (1992, 1995), Tegene & Kuchler (1994) and Wang & Bessler (2004) provided empirical evidence to suggest that ECMs outperform VARs in levels, particularly over longer forecast horizons. Also Shoesmith (1995), and later Villani (2001), showed how Litterman's (1986) Bayesian approach can improve forecasting with cointegrated VARs. Reimers (1997) studied the forecasting performance of seasonally cointegrated vector time series processes using an ECM in fourth differences. Poskitt (2003) discussed the specification of cointegrated VARMA systems. Chevillon & Hendry (2005) analyzed the relation between direct multi-step estimation of stationary and non-stationary VARs and forecast accuracy.

4 Seasonality

The oldest approach to handling seasonality in time series is to extract it using a seasonal decomposition procedure such as the X-11 method. Over the past 25 years, the X-11 method and its variants have been studied extensively.

One line of research has considered the effect of using forecasting as part of the seasonal decomposition method. For example, Dagum (1982) and Huot et al. (1986) looked at the use of forecasting in X-11-ARIMA to reduce the size of revisions in the seasonal adjustment of data, and Pfeffermann et al. (1995) explored the effect of the forecasts on the variance of the trend and seasonally adjusted values.

Quenneville et al. (2003) took a different perspective and looked at forecasts implied by the asymmetric moving average filters in the X-11 method and its variants.

A third approach has been to look at the effectiveness of forecasting using seasonally adjusted data obtained from a seasonal decomposition method. Miller & Williams (2003, 2004) showed that greater forecasting accuracy is obtained by shrinking the seasonal component towards zero. The commentaries on the latter paper (Findley et al., 2004; Ladiray & Quenneville, 2004; Hyndman, 2004; Koehler, 2004; and Ord, 2004) gave several suggestions regarding implementation of this idea.

In addition to work on the X-11 method and its variants, there have also been several new methods for seasonal adjustment developed, the most important being the model based approach of TRAMO-SEATS (Gómez & Maravall, 2001) and the nonparametric method STL (Cleveland et al., 1990). Another proposal has been to use sinusoidal models (Simmons, 1990).

When forecasting several similar series, Withycombe (1989) showed that it can be more efficient to estimate a combined seasonal component from the group of series, rather than individual seasonal patterns. Bunn & Vassilopoulos (1993) demonstrated how to use clustering to form appropriate groups for this situation, and Bunn & Vassilopoulos (1999) introduced some improved estimators for the group seasonal indices.

Twenty five years ago, unit root tests had only recently been invented, and seasonal unit root tests were yet to appear. Subsequently, there has been considerable work done on the use and implementation of seasonal unit root tests including Hylleberg & Pagan (1997), Taylor (1997) and Franses & Koehler (1998). Paap et al. (1997) and Clements & Hendry (1997) studied the forecast performance of models with unit roots, especially in the context of level shifts.

Some authors have cautioned against the widespread use of standard seasonal unit root models for economic time series. Osborn (1990) argued that deterministic seasonal components are more common in economic series than stochastic seasonality. Franses & Romijn (1993) suggested that seasonal roots in periodic models result in better forecasts. This suggestion was evaluated by Wells (1997), Herwartz (1997) and Novales & de Fruto (1997).

Several papers have compared various seasonal models empirically. Chen (1997) explored the robustness properties of a structural model, a regression model with seasonal dummies, an ARIMA model, and Holt-Winters' method, and found that the latter two yield forecasts that are relatively robust to model misspecification. Noakes et al. (1985), Albertson & Aylen (1996), Kulendran & King (1997) and Franses & van Dijk (2005) each compared the forecast performance of several seasonal models applied to real data.

5 State space and structural models and the Kalman filter

At the start of the 1980s, state space models were only beginning to be used by statisticians for forecasting time series, although the ideas had been present in the engineering literature since Kalman's (1960) ground-breaking work. State space models provide a unifying framework in which any linear time series model can be written. The key forecasting contribution of Kalman (1960) was to give a recursive algorithm (known as the Kalman filter) for computing forecasts. Statisticians became interested in state space models when Schweppe (1965) showed that the Kalman filter provides an efficient algorithm for computing the one-step-ahead prediction errors and associated variances need to produce the likelihood function.

A particular class of state space models, known as "structural models" or "dynamic linear models" (DLM), was introduced by Harrison & Stevens (1976), who also proposed a Bayesian approach to estimation. These models bear many similarities with exponential smoothing methods, but have multiple sources of random error. In particular, the "basic structural model" (BSM) is similar to Holt-Winters' method for seasonal data and includes a level, trend and seasonal component. Harvey (1984, 1989) extended the class of structural models and followed a non-Bayesian approach to estimation.

Ray (1989) discussed convergence rates for the linear growth structural model and showed that the initial states (usually chosen subjectively) have a non-negligible impact on forecasts. Harvey & Snyder (1990) proposed some continuous-time structural models for use in forecasting lead time demand for inventory control. Proietti (2000) discussed several variations on the BSM and compared their properties and evaluated the resulting forecasts.

Another class of state space models, known as "balanced state space models", has been used primarily for forecasting macroeconomic time series. Mittnik (1990) provided a survey of this class of models, and Vinod & Basu (1995) obtained forecasts of consumption, income and interest rates using balanced state space models. These models have only one source of random error and subsume various other time series models including ARMAX models, ARMA models and rational distributed lag models. A related class of state space models are the "single source of error" models that underly exponential smoothing methods; these are discussed in Section 2.

As well as these methodological developments, there have been several papers proposing innovative state space models to solve practical forecasting problems. These include Coomes (1992) who used a state space model to forecast jobs by industry for local regions, and Patterson (1995) who used a state space approach for forecasting real personal disposable income.

6 Nonlinear models

6.1 Preamble

Relative to linear time series study, the development of nonlinear time series analysis and forecasting is still in its infancy. The beginning of nonlinear time series has been attributed to Volterra (1930). He showed that any continuous nonlinear function in t could be approximated by a finite Volterra series. Wiener (1958) became interested in the ideas of functional series representation, and further developed the existing material. Although the probabilistic properties of these models have been studied extensively, the problems of parameter estimation, model fitting, and forecasting, have been neglected for a long time. This neglect can largely be attributed to the complexity of the proposed Wiener model, and its simplified forms like the bilinear model (Poskitt & Tremayne, 1986). At the time, fitting these models led to what were insurmountable computational difficulties.

Although linearity is a useful assumption and a powerful tool in many areas, it became increasingly clear in the late 1970s and early 1980s that linear models are insufficient in many real applications. For example, sustained animal population size cycles (the well-known Canadian lynx data), sustained solar cycles (annual sunspot numbers), energy flow and amplitude-frequency relations were found not to be suitable for linear models. Accelerated by practical demands, several useful nonlinear time series models were proposed in this same period. De Gooijer & Kumar (1992) provided an overview of the developments in this area to the beginning of the 1990s. These authors argued that the evidence for the superior forecasting performance of nonlinear models is patchy. One factor that has probably retarded the widespread reporting of nonlinear forecasts is that up to that time it was not possible to obtain closed-form analytic expressions for multi-step-ahead forecasts.

The monograph by Granger & Teräsvirta (1993) has boosted new developments in estimating, evaluating, and selecting among nonlinear forecasting models for economic and financial time series. A good overview of the current state-of-the-art is *IJF* Special Issue 20:2 (2004). In their introductory paper Clements et al. (2004) outlined a variety of topics for future research. They concluded that "… the day is still long off when simple, reliable and easy to use non-linear model specification, estimation and forecasting procedures will be readily available".

6.2 Regime switching models

The class of (self-exciting) threshold AR (SETAR) models was introduced by Tong (1978). These models, which are piecewise linear models in their most basic form, have attracted some attention in the *IJF*. Clements & Smith (1997) compared a number of methods for obtaining multi-step-ahead forecasts for univariate discrete-time SETAR models. They concluded that forecasts made using Monte Carlo simulation are satisfactory in cases were it is known that the disturbances in the SETAR model come from a symmetric distribution. Otherwise the Bootstrap method is to be preferred. Similar results were reported by De Gooijer & Vidiella-i-Anguera (2004) for threshold VAR models. Brockwell & Hyndman (1992) obtained one-step-ahead forecasts for univariate continuous-time threshold AR models (CTAR). Since the calculation of multi-step-ahead forecasts from CTAR models involves complicated higher dimensional integration, the practical use of CTARs is limited. The out-of-sample forecast performance of various variants of SETAR models relative to linear models has been the subject of several *IJF* papers, including Astatkie et al. (1997), Boero & Marrocu (2004) and Enders & Falk (1998).

One drawback of the SETAR model is that the dynamics change discontinuously from one regime to the other. In constrast, a smooth transition AR (STAR) model allows for a more gradual transition between the different regimes. Sarantis (2001) found evidence that STAR-type models can improve upon linear AR and random walk models in forecasting stock prices at both short term and medium term horizons. Interestingly, the recent study by Bradley & Jansen (2004) seems to refute Sarantis' conclusion.

Franses et al. (2004) proposed a threshold AR(1) model that allows for plausible inference about the specific values of the parameters. The key idea is that the values of the AR parameter depend on a leading indicator variable. The resulting model outperforms other time-varying nonlinear models, including the Markow regime-switching model, in terms of forecasting.

6.3 Neural nets

The artificial neural network (ANN) can be useful for nonlinear processes that have an unknown functional relationship and as a result are difficult to fit (Darbellay & Slama, 2000). The main idea with ANNs is that inputs, or dependent variables, get filtered through one or more hidden layers each of which consist of hidden units, or nodes, before they reach the output variable. Next the intermediate output is related to the final output. Various other nonlinear models are specific versions of ANNs, where more structure is imposed.

One major application area of ANNs is forecasting; see Zhang et al. (1998) for a good survey of the literature. Numerous studies outside the *IJF* have documented the successes of ANNs in forecasting financial data. However, in two editorials in this *Journal*, Chatfield (1993, 1995) questioned whether ANNs had been oversold as a miracle forecasting technique. This was followed by several papers documenting that naïve financial models such as the random walk out-perform ANNs (see, e.g., Church & Curram, 1996; Callen et al., 1996; Gorr et al., 1994; Tkacz, 2001).

Gorr (1994) and Hill et al. (1994) suggested that future research should investigate and better define the borderline between where ANNs and "traditional" techniques out-perform one other. That theme is explored by several authors. Hill et al. (1994) noticed that ANNs are likely to work best for high frequency financial data and Balkin & Ord (2000) also stressed the importance of a long time series to ensure optimal results from training ANNs. Qi (2001) pointed out that ANNs are more likely to out-perform other methods when the input data is kept as current as possible, using recursive modelling (see also Olson & Mossman, 2003).

A general problem with nonlinear models is the "curse of model complexity and model over-parametrization". If parsimony is considered to be really important, then it interesting to compare the out-of-sample forecasting performance of linear versus nonlinear models, using a wide variety of different model selection criteria. This issue was considered in quite some depth by Swanson & White (1997). Their results suggested that a single hidden layer 'feedforward' ANN model, which has been by far the most popular in time series econometrics, offers a useful and flexible alternative to fixed specification linear models, particularly at forecast horizons greater than one-step-ahead. However, in contrast to Swanson & White, Heravi et al. (2004) found that linear models produce more accurate forecasts of monthly seasonally unadjusted European industrial production series than ANN models. Ghiassa et al. (2005) presented a dynamic ANN and compared its forecasting performance against the traditional ANN and ARIMA models.

6.4 Deterministic versus stochastic dynamics

The possibility that nonlinearities in high-frequency financial data (e.g., hourly returns) are produced by a low-dimensional deterministic chaotic process has been the subject of a few studies published in the *IJF*. Cecen & Erkal (1996) showed that it is not possible to exploit deterministic nonlinear dependence in daily spot rates in order to improve short-term forecasting. Lisi & Medio (1997) reconstructed the state space for a number of monthly exchange rates and, using a local linear method, approximated the dynamics of the system on that space. One-step-ahead out-of-sample forecasting showed that their method outperforms a random walk model. A similar study was performed by Cao & Soofi (1999).

6.5 Miscellaneous

A host of other, often less well known, nonlinear models have been used for forecasting purposes. For instance, Ludlow & Enders (2000) adopted Fourier coefficients to approximate the various types of nonlinearities present in time series data. Herwartz (2001) extended the linear vector ECM to allow for asymmetries. Dahl & Hylleberg (2004) compared Hamilton's (2001) flexible nonlinear regression model, ANNs, and two versions of the projection pursuit regression model. Time-varying AR models are included in a comparative study by Marcellino (2004). The nonparametric, nearest-neighbour method was applied by Fernández-Rodríguez et al. (1999).

7 Long memory

When the integration parameter d in an ARIMA process is fractional and greater than zero, the process exhibits long memory in the sense that observations a long time-span apart have non-negligible dependence. Stationary long-memory models ($0 < d < 0.5$), also termed fractionally differenced ARMA (FARMA) or fractionally integrated ARMA (ARFIMA) models, have been considered by workers in many fields; see Granger & Joyeux (1980) for an introduction. One motivation for these studies is that many empirical time series have a sample autocorrelation function which declines at a slower rate than for an ARIMA model with finite orders and integer d.

The forecasting potential of fitted FARMA/ARFIMA models, as opposed to forecast results obtained from other time series models, has been a topic of various *IJF* papers and a special issue (2002, 18:2). Ray (1993) undertook such a comparison between seasonal FARMA/ARFIMA models and standard (non-fractional) seasonal ARIMA models. The results show that higher order AR models are capable of forecasting the longer term well when compared with ARFIMA models. Following Ray (1993), Smith & Yadav (1994) investigated the cost of assuming a unit difference when a series is only fractionally integrated with $d \neq 1$. Over-differencing a series will produce a loss in forecasting performance one-step-ahead, with only a limited loss thereafter. By contrast, under-differencing a series is more costly with larger potential losses from fitting a mis-specified AR model at all forecast horizons. This issue is further explored by Andersson (2000) who showed that misspecification strongly affects the estimated memory of the ARFIMA model, using a rule which is similar to the test of Öller (1985). Man (2003) argued that a suitably adapted ARMA(2,2) model can produce short-term forecasts that are competitive with estimated ARFIMA models. Multi-step ahead forecasts of long memory models have been developed by Hurvich (2002), and compared by Bhansali & Koskoska (2002).

Many extensions of ARIFMA models and a comparison of their relative forecasting performance have been explored. For instance, Franses & Ooms (1997) proposed the so-called periodic ARFIMA$(0, d, 0)$ model where d can vary with the seasonality parameter. Ravishankar & Ray (2002) considered the estimation and forecasting of multivariate ARFIMA models. Baillie & Chung (2002) discussed the use of linear trend-stationary ARFIMA models, while the paper by Beran et al. (2002) extended this model to allow for nonlinear trends. Souza & Smith (2002) investigated the effect of different sampling rates, such as monthly versus quarterly data, on estimates of the long-memory parameter d. In a similar vein, Souza & Smith (2004) looked at the effects of temporal aggregation on estimates and forecasts of ARFIMA processes.

8 ARCH/GARCH models

A key feature of financial time series is that large (small) absolute returns tend to be followed by large (small) absolute returns, that is, there are periods which display high (low) volatility. The class of generalized autoregressive conditionally heteroscedastic (GARCH) models introduced by Engle (1982) are considered extremely useful for modelling and forecasting movements in asset return volatilities over time; see Bollerslev et al. (1992) for a comprehensive review. Sabbatini & Linton (1998) showed that the simple (linear) GARCH(1,1) model is a good parametrization for the daily returns on the Swiss market index. However, the quality of the out-of-sample forecasts suggests that this result should be taken with caution. Franses & Ghijsels (1999) stressed that this feature can be due to neglected additive outliers (AO). They noted that GARCH models for AO-corrected returns results in improved forecasts of stock market volatility. At the estimation level, Brooks et al. (2001) argued that standard econometric software packages can produce widely varying results. Clearly, this may have some impact on the forecasting accuracy of GARCH models.

Using two daily exchange rates series, Galbraith & Kisinbay (2005) compared the forecast content functions both from the standard GARCH model and from a fractionally integrated GARCH model, at horizons of up to 30 trading days.

Empirically, returns and conditional variance of next period's returns are negatively correlated. That is, negative (positive) returns are generally associated with upward (downward) revisions of the conditional volatility. This phenomenon is often referred to as asymmetric volatility in the literature; see, e.g., Engle and Ng (1993). Various nonlinear GARCH have been developed by researchers to capture asymmetric volatility; see, e.g. Hentschel (1995) and Pagan (1996) for overviews. Awartani & Corradi (2005) investigated the impact of asymmetries on the out of sample forecast ability of different GARCH models, at various horizons.

9 Count data forecasting

Count data occur frequently in business and industry, especially in inventory data where they are often called "intermittent demand data". Consequently, it is surprising that so little work has been done on forecasting count data. Some work has been done on ad hoc methods for forecasting count data, but few papers have appeared on forecasting count time series using stochastic models.

Most work on count forecasting is based on Croston (1972) who proposed using SES to independently forecast the nonzero values of a series and the time between non-zero values. Willemain et al. (1994) compared Croston's method to SES and found that Croston's method was more robust, although these results were based on MAPEs which are often undefined for count data. The conditions under which Croston's method does better than SES are discussed in Johnston & Boylan (1996). Willemain et al. (2004) proposed a bootstrap procedure for intermittent demand data which was more accurate than either SES or Croston's method on the nine series evaluated.

Evaluating count forecasts raises difficulties due to the presence of zeros in the observed data. Syntetos & Boylan (2005) proposed using the Relative Mean Absolute Error (see Section 10), while Willemain et al. (2004) recommended using the probability-integral transform method of Diebold et al. (1998).

Grunwald et al. (2000) surveyed many of the stochastic models for count time series, using simple first-order autoregression as a unifying framework for the various approaches. One possible model, explored by Brännäs (1995), assumes the series follows a Poisson distribution with a mean that depends on an unobserved and autocorrelated process. An alternative integer-valued MA model was used by Brännäs et al. (2002) to forecast occupancy levels in Swedish hotels.

The forecast distribution can be obtained by simulation using any of these stochastic models, but how to summarize the distribution is not obvious. Freeland & McCabe (2004) proposed using the median of the forecast distribution, and gave a method for computing confidence intervals for the entire forecast distribution in the case of integer-valued autoregressive

(INAR) models of order 1. McCabe & Martin (2005) further extended these ideas by presenting a Bayesian methodology for forecasting from the INAR class of models.

10 Forecast evaluation and accuracy measures

A bewildering array of accuracy measures have been used to evaluate the performance of forecasting methods. Some of them are listed in the early survey paper of Mahmoud (1984). We first define the most common measures.

Let Y_t denote the observation at time t and F_t denote the forecast of Y_t. Then define the forecast error $e_t = Y_t - F_t$ and the percentage error as $p_t = 100e_t/Y_t$. An alternative way of scaling is to divide each error by the error obtained with another standard method of forecasting. Let $r_t = e_t/e_t^*$ denote the relative error where e_t^* is the forecast error obtained from the base method. Usually, the base method is the "naïve method" where F_t is equal to the last observation. We use the notation mean(x_t) to denote the sample mean of $\{x_t\}$ over the period of interest (or over the series of interest). Analogously, we use median(x_t) for the sample median and gmean(x_t) for the geometric mean. The mostly commonly used methods are defined in Table 2 below. Here, MAE$_b$ (MSE$_b$) refers to the MAE (MSE) obtained from the base method.

Note that Armstrong & Collopy (1992) referred to RelMAE as CumRAE, and that RelRMSE is also known as Thiel's U statistic (Theil, 1966, Chapter 2) and is sometimes called $U2$. In addition to these, the average ranking (AR) of a method relative to all other methods considered, has sometimes been used.

MSE	Mean Square Error	$= \text{mean}(e_t^2)$
RMSE	Root Mean Square Error	$= \sqrt{\text{MSE}}$
MAE	Mean Absolute Error	$= \text{mean}(\|e_t\|)$
MdAE	Median Absolute Error	$= \text{median}(\|e_t\|)$
MAPE	Mean Absolute Percentage Error	$= \text{mean}(\|p_t\|)$
MdAPE	Median Absolute Percentage Error	$= \text{median}(\|p_t\|)$
sMAPE	Symmetric Mean Absolute Percentage Error	$= \text{mean}(2\|Y_t - F_t\|/(Y_t + F_t))$
sMdAPE	Symmetric Median Absolute Percentage Error	$= \text{median}(2\|Y_t - F_t\|/(Y_t + F_t))$
MRAE	Mean Relative Absolute Error	$= \text{mean}(\|r_t\|)$
MdRAE	Median Relative Absolute Error	$= \text{median}(\|r_t\|)$
GMRAE	Geometric Mean Relative Absolute Error	$= \text{gmean}(\|r_t\|)$
RelMAE	Relative Mean Absolute Error	$= \text{MAE}/\text{MAE}_b$.
RelRMSE	Relative Root Mean Squared Error	$= \text{RMSE}/\text{RMSE}_b$.
LMR	Log Mean Squared Error Ratio	$= \log(\text{RelMSE})$
PB	Percentage Better	$= 100\,\text{mean}(I\{r_t < 1\})$
PB(MAE)	Percentage Better (MAE)	$= 100\,\text{mean}(I\{\text{MAE} < \text{MAE}_b\})$
PB(MSE)	Percentage Better (MSE)	$= 100\,\text{mean}(I\{\text{MSE} < \text{MSE}_b\})$

Table 2: *Commonly used forecast accuracy measures.*

The evolution of measures of forecast accuracy and evaluation can be seen through the measures used to evaluate methods in the major comparative studies that have been undertaken. In the original M-competition (Makridakis et al., 1982), measures used included the MAPE, MSE, AR, MdAPE and PB. However, as Chatfield (1988) and Armstrong & Collopy (1992) pointed out, the MSE is not appropriate for comparison between series as it is scale dependent. Fildes & Makridakis (1988) contained further discussion on this point. The MAPE also has problems when the series has values close to (or equal to) zero, as noted by Makridakis et al. (1998, p. 45). Excessively large (or infinite) MAPEs were avoided in the M-competitions by only including data that were positive. However, this is an artificial solution that is impossible to apply in practical situations.

In 1992, one issue of *IJF* carried two articles and several commentaries on forecast evaluation measures. Armstrong & Collopy (1992) recommended the use of relative absolute errors, especially the GMRAE and MdRAE, despite the fact that relative errors have infinite variance and undefined mean. They recommended "winsorizing" to trim extreme values which will partially overcome these problems, but which adds some complexity to the calculation and a level of arbitrariness as the amount of trimming must be specified. Fildes (1992) also preferred the GMRAE although he expressed it in an equivalent (but unnecessarily confusing) form as the square root of the geometric mean of squared relative errors. This equivalence does not seem to have been noticed by any of the discussants in the commentaries of Ahlburg et al. (1992).

The study of Fildes et al. (1998), which looked at forecasting telecommunications data, used MAPE, MdAPE, PB, AR, GMRAE and MdRAE, taking into account some of the criticism of the methods used for the M-competition.

The M3-competition (Makridakis & Hibon, 2000) used three different measures of accuracy: MdRAE, sMAPE and sMdAPE. The "symmetric" measures were proposed by Makridakis (1993) in response to the observation that the MAPE and MdAPE have the disadvantage that they put a heavier penalty on positive errors than on negative errors. However, these measures are not as "symmetric" as their name suggests. For the same value of Y_t, the value of $2|Y_t - F_t|/(Y_t + F_t)$ has a heavier penalty when forecasts are high compared to when forecasts are low. See Goodwin & Lawton (1999) and Koehler (2001) for further discussion on this point.

Notably, none of the major comparative studies have used relative measures (as distinct from measures using relative errors) such as RelMAE or LMR. The latter was proposed by Thompson (1990) who argued for its use based on its good statistical properties. It was applied to the M-competition data in Thompson (1991).

Apart from Thompson (1990), there has been very little theoretical work on the statistical properties of these measures. One exception is Wun & Pearn (1991) who looked at the statistical properties of MAE.

A novel alternative measure of accuracy is "time distance" which was considered by Granger & Jeon (2003a,b). In this measure, the leading and lagging properties of a forecast are also captured. Again, this measure has not been used in any major comparative study.

A parallel line of research has looked at statistical tests to compare forecasting methods. An early contribution was Flores (1989). The best known approach to testing differences between the accuracy of forecast methods is the Diebold-Mariano (1995) test. A size-corrected modification of this test was proposed by Harvey et al. (1997). McCracken (2004) looked at the effect of parameter estimation on such tests and provided a new method for adjusting for parameter estimation error.

Another problem in forecast evaluation, and more serious than parameter estimation error, is "data sharing"—the use of the same data for many different forecasting methods. Sullivan et al. (2003) proposed a bootstrap procedure designed to overcome the resulting distortion of statistical inference.

11 Combining

Combining, mixing, or pooling quantitative[4] forecasts obtained from very different time series methods and different sources of information has been studied for the past three decades. Important early contributions in this area were made by Bates & Granger (1969), Newbold & Granger (1974) and Winkler & Makridakis (1983). Compelling evidence on the relative efficiency of combined forecasts, usually defined in terms of forecast error variances, was summarized by Clemen (1989) in a comprehensive bibliography review.

Numerous methods for selecting the combining weights have been proposed. The simple average is the most-widely used combining method (see Clemen's review, and Bunn, 1985), but the method does not utilize past information regarding the precision of the forecasts or the dependence among the forecasts. Another simple method is a linear mixture of the individual forecasts with combining weights determined by OLS (assuming unbiasedness) from the matrix of past forecasts and the vector of past observations (Granger & Ramanathan, 1984). However, the OLS estimates of the weights are inefficient due to the possible presence of serial correlation in the combined forecast errors. Aksu & Gunter (1992) and Gunter (1992) investigated this problem in some detail. They recommended the use of OLS combination forecasts with the weights restricted to sum to unity.

Rather than using fixed weights, Deutsch et al. (1994) allowed them to change through time using switching regime models and STAR models. Another time-dependent weighting scheme was proposed by Fiordaliso (1998), who used a fuzzy system to combine a set of individual forecasts in a nonlinear way. Diebold & Pauly (1990) used Bayesian shrinkage techniques to allow the incorporation of prior information into the estimation of combining weights. Combining forecasts from very *similar* models, with weights sequentially updated, was considered by Zou & Yang (2004).

Combining weights determined from time-invariant methods can lead to relatively poor forecasts if nonstationarity among component forecasts occurs. Miller et al. (1992) examined the

[4]See Kamstra & Kennedy (1998) for a computationally-convenient method of combining qualitative forecasts.

effect of 'location-shift' nonstationarity on a range of forecast combination methods. Tentatively they concluded that the simple average beats more complex combination devices; see also Hendry & Clements (2002) for more recent results. The related topic of combining forecasts from linear and some nonlinear time series models, with OLS weights as well as weights determined by a time-varying method, was addressed by Terui & van Dijk (2002).

The shape of the combined forecast error distribution and the corresponding stochastic behaviour was studied by de Menezes & Bunn (1998) and Taylor & Bunn (1999). For non-normal forecast error distributions skewness emerges as a relevant criterion for specifying the method of combination. Some insights into why competing forecasts may be fruitfully combined to produce a forecast superior to individual forecasts were provided by Fang (2003), using forecast encompassing tests. Hibon & Evgeniou (2005) proposed a criterion to select among forecasts and their combinations.

12 Prediction intervals and densities

The use of prediction intervals, and more recently prediction densities, has become much more common over the past twenty five years as practitioners have come to understand the limitations of point forecasts. Unfortunately, there is still some confusion on terminology with many authors using "confidence interval" instead of "prediction interval". A confidence interval is for a model parameter whereas a prediction interval is for a random variable. Almost always, forecasters will want prediction intervals—intervals which contain the true values of future observations with specified probability.

Most prediction intervals are based on an underlying stochastic model. Consequently, there has been a large amount of work on formulating appropriate stochastic models underlying some common forecasting procedures (see, e.g., Section 2 on Exponential Smoothing).

The link between prediction interval formulae and the model from which they are derived has not always been correctly observed. For example, the prediction interval appropriate for a random walk model was applied by Makridakis & Hibon (1987) and Lefrançois (1989) to forecasts obtained from many other methods. This problem was noted by Koehler (1990) and Chatfield & Koehler (1991).

With most model-based prediction intervals for time series, the uncertainty associated with model selection and parameter estimation is not accounted for. Consequently, the intervals are too narrow. There has been considerable research on how to make model-based prediction intervals have more realistic coverage. A series of papers on using the bootstrap to compute prediction intervals for an AR model has appeared beginning with Masarotto (1990), and including Grigoletto (1998), Clements & Taylor (2001) and Kim (2004). Similar procedures for other models have also been considered including ARIMA models (Pascual et al., 2001, 2005), VAR (Kim, 1999), ARCH (Reeves, 2005) and regression (Lam & Veall, 2002).

When the forecast error distribution is non-normal, finding the entire forecast density is useful as a single interval may no longer provide an adequate summary of the expected future. A review of density forecasting is provided by Tay & Wallis (2000), along with several other articles in the same special issue of the *JoF*. Summarizing, a density forecast has been the subject of some interesting proposals including "fan charts" (Wallis, 1999) and "highest density regions" (Hyndman, 1995).

As prediction intervals and forecast densities have become more commonly used, attention has turned to their evaluation and testing. Wallis (2003) proposed chi-squared tests for both intervals and densities, and Clements & Smith (2002) discussed some simple but powerful tests when evaluating multivariate forecast densities.

13 A look to the future

In the preceding sections we looked back at the time series forecasting history of the *IJF*, in the hope that the past may shed light on the present. But a silver anniversary is also a good time to look ahead. However, before doing so, it is interesting to summarize the proposals for research in time series forecasting identified in a set of related papers by Ord, Cogger and Chatfield published in this Journal more than 15 years ago.

Ord (1988) noted that not much work has been done on multiple time series models, including multivariate exponential smoothing. He also indicated the need for deeper research in forecasting methods based on nonlinear models. While many aspects of both areas have been investigated in the *IJF*, they merit continued research. For instance, there is still no clear consensus that forecasts from nonlinear models substantively outperform those from linear models; see, e.g., Stock & Watson (1999).

Given the frequent misuse of methods based on linear models with Gaussian i.i.d. distributed errors, Cogger (1988) argued that new developments in the area of 'robust' statistical methods should receive more attention within the time series forecasting community. A robust procedure is expected to work well when there are outliers or location shifts in the data that are hard to detect. Robust statistics can be based on both parametric and nonparametric methods. An example of the latter is the Koenker-Bassett (1978) concept of regression quantiles investigated by Cogger. In retrospect, it is fair to say that the *IJF* has not attracted many papers in this area. The reason may be the high level of technicality of many robust methods. Nevertheless, outside the *IJF* some progress has been made. For instance, Yao & Tong (1995) proposed the concept of conditional percentile prediction interval. Its width is no longer a constant, as in the case of linear models, but may vary with respect to the position in the state space from which forecasts are being made; see also De Gooijer & Gannoun (2000) and Polonik & Yao (2000). Clearly, the area of robustifying forecasts and forecast intervals requires further research; see item 14.13 on Armstrong's (2001) list of 23 principles that are in great need of research.

Chatfield (1988) stressed the need for future research on developing multivariate methods with an emphasis on making them more of a practical proposition. The works on state space, Kalman filtering, and discrete/continuous time structural models by Harvey (1989), West &

Harrison (1989, 1997) and Durbin & Koopman (2001) have had an impact on the time series literature. However, as noted earlier, forecasting applications of the state space framework using the Kalman filter has been rather limited in the *IJF*. In that sense, it is perhaps not too surprising that even today, some textbook authors do not seem to realize that the Kalman filter can, for example, track a nonstationary process stably.

In the 1990s some *IJF* authors tried to identify new important research topics. Both De Gooijer (1990) and Clements (2003) in two Editorials, and Ord as a part of a discussion paper by Dawes et al. (1994), suggested more work on combining forecasts. Although the topic received a fair amount of attention (see Section 11) there are still several open questions. For instance, what is the "best" combining method for linear and nonlinear models, and what prediction interval can be put around the combined forecast? See also Armstrong (2001, items 12.5–12.7). Other topics suggested by Ord include the need to develop model selection procedures that make effective use of both data and prior knowledge, and the need to specify objectives for forecasts and develop forecasting systems that address those objectives.

There are many distinct directions in which the field of time series forecasting could proceed, some of which have already been noted above. If we look at the areas underrepresented in the *IJF* the following potential research themes emerge:

- Conditional quantiles. Relatively little attention has been paid to the use of univariate and multivariate conditional quantiles as a tool for accurate forecasts in economics. One important area of application is in estimating risk management tools such as Value-at-Risk. Recently, Engle & Manganelli (2004) made a start in this direction, proposing a conditional value at risk model.
- Panel data models/longitudinal data analysis. In many panels analyzed in the past the time series dimension t has been small whilst the cross-section dimension n is large. However, nowadays in many applied areas such as marketing large datasets can be easily collected with n and t both large. Extracting features from megapanels of longitudinal data is well-known under the name "functional data analysis"; see, e.g., Ramsay & Silverman (1997, 2005). Yet, the problem of making multi-step ahead forecasts based on functional data is still open for both theoretical and applied research.
- Count forecasting. Forecasting time series of counts is an area still in its infancy, despite the prevalence of these data in business and industry. There are many unresolved theoretical and practical problems associated with count forecasting, and we expect much productive research in this area over the next few years.
- Computationally intensive methods for large data sets. While neural networks have been used in forecasting for more than a decade now, there are many outstanding issues associated with their use and implementation, including when they are likely to out-perform other methods. Other methods involving heavy computation (e.g., bagging and boosting) are even less understood in the forecasting context. With the availability of very large data sets and high powered computers, we expect this to be an important area of research in the coming years.

References

Exponential smoothing

Archibald, B.C. (1990), “Parameter space of the Holt-Winters’ model”, *International Journal of Forecasting*, 6, 199–209.

Archibald, B.C. and A.B. Koehler (2003), “Normalization of seasonal factors in Winters’ methods”, *International Journal of Forecasting*, 19, 143–148.

Assimakopoulos, V. and K. Nikolopoulos (2000), ‘The theta model: A decomposition approach to forecasting’, *International Journal of Forecasting*, 16, 521–530.

Bartolomei, S.M. and A.L. Sweet (1989), ‘A note on a comparison of exponential smoothing methods for forecasting seasonal series’, *International Journal of Forecasting*, 5, 111–116.

Brown, R.G. (1959), *Statistical forecasting for inventory control*, New York: McGraw-Hill.

Brown, R.G. (1963), *Smoothing, forecasting and prediction of discrete time series*, Englewood Cliffs, NJ: Prentice-Hall.

Carreno, J. and Madinaveitia, J. (1990), ‘A modification of time series forecasting methods for handling announced price increases’, *International Journal of Forecasting*, 6, 479–484.

Chatfield, C., A.B. Koehler, J.K. Ord and Snyder, R.D. (2001), ‘A new look at models for exponential smoothing’, *The Statistician*, 50, 147–159.

Chatfield, C. and Yar, M. (1991), ‘Prediction intervals for multiplicative Holt-Winters’, *International Journal of Forecasting*, 7, 31–37.

Collopy, F. and J.S. Armstrong (1992), ‘Rule-based forecasting: development and validation of an expert systems approach to combining time series extrapolations’, *Management Science*, 38, 1394–1414.

Gardner, Jr., E.S. (1985), ‘Exponential smoothing: the state of the art’, *Journal of Forecasting*, 4, 1–37 (with commentaries).

Gardner, Jr., E.S. (1993), ‘Forecasting the failure of component parts in computer systems: A case study’, *International Journal of Forecasting*, 9, 245–253.

Gardner, Jr., E.S. and E. McKenzie (1988), ‘Model identification in exponential smoothing’, *Journal of the Operational Research Society*, 39, 863–867.

Grubb, H. and A. Masa (2001), ‘Long lead-time forecasting of UK air passengers by Holt-Winters methods with damped trend’, *International Journal of Forecasting*, 17, 71–82.

Holt, C.C. (1957), ‘Forecasting seasonals and trends by exponentially weighted averages’, O.N.R. Memorandum 52/1957, Carnegie Institute of Technology.

Holt, C.C. (2004), ‘Forecasting seasonals and trends by exponentially weighted moving averages’ (with discussion), *International Journal of Forecasting*, 20, 5–13.

Hyndman, R.J. (2001), ‘It’s time to move from what to why’, *International Journal of Forecasting*, 17, 567–570.

Hyndman, R.J. and B. Billah (2003), ‘Unmasking the Theta method’, *International Journal of Forecasting*, 19, 287–290.

Hyndman, R.J., A.B. Koehler, J.K. Ord and R.D. Snyder (2005), ‘Prediction intervals for exponential smoothing state space models’, *Journal of Forecasting*, 24, 17–37.

Hyndman, R.J., A.B. Koehler, R.D. Snyder and S. Grose (2002), ‘A state space framework for automatic forecasting using exponential smoothing methods’, *International Journal of Forecasting*, 18, 439–454.

Johnston, F.R. and P.J. Harrison (1986), 'The variance of lead-time demand', *Journal of Operational Research Society*, 37, 303–308.

Koehler, A.B., R.D. Snyder and J.K. Ord (2001), 'Forecasting models and prediction intervals for the multiplicative Holt-Winters method', *International Journal of Forecasting*, 17, 269–286.

Lawton, R. (1998), 'How should additive Holt-Winters estimates be corrected?', *International Journal of Forecasting*, 14, 393–403.

Ledolter, J. and B. Abraham (1984), 'Some comments on the initialization of exponential smoothing', *Journal of Forecasting*, 3, 79–84.

Makridakis, S. and M. Hibon (1991), 'Exponential smoothing: The effect of initial values and loss function on post-sample forecasting accuracy', *International Journal of Forecasting*, 7, 317–330.

McClain, J.G. (1988), 'Dominant tracking signals', *International Journal of Forecasting*, 4, 563–572.

McKenzie, E. (1984), 'General exponential smoothing and the equivalent ARMA process', *Journal of Forecasting*, 3, 333–344.

McKenzie, E. (1986), "Error analysis for Winters' additive seasonal forecasting system", *International Journal of Forecasting*, 2, 373–382.

Miller, T. and M. Liberatore (1993), 'Seasonal exponential smoothing with damped trends. An application for production planning', *International Journal of Forecasting*, 9, 509–515.

Newbold, P. and T. Bos (1989), 'On exponential smoothing and the assumption of deterministic trend plus white noise data-generating models', *International Journal of Forecasting*, 5, 523–527.

Ord, J.K., A.B. Koehler and R.D. Snyder (1997), 'Estimation and prediction for a class of dynamic nonlinear statistical models', *Journal of American Statistical Association*, 92, 1621–1629.

Pegels, C.C. (1969), 'Exponential smoothing: some new variations', *Management Science*, 12, 311–315.

Pfeffermann, D. and J. Allon (1989), 'Multivariate exponential smoothing: methods and practice', *International Journal of Forecasting*, 5, 83–98.

Roberts, S.A. (1982), 'A general class of Holt-Winters type forecasting models', *Management Science*, 28, 808–820.

Rosas, A.L. and V.M. Guerrero (1994), 'Restricted forecasts using exponential smoothing techniques', *International Journal of Forecasting*, 10, 515–527.

Satchell, S. and A. Timmermann (1995), 'On the optimality of adaptive expectations: Muth revisited', *International Journal of Forecasting*, 11, 407–416.

Snyder, R.D. (1985), 'Recursive estimation of dynamic linear statistical models', *Journal of the Royal Statistical Society (B)*, 47, 272–276.

Sweet, A.L. (1985), 'Computing the variance of the forecast error for the Holt-Winters seasonal models', *Journal of Forecasting*, 4, 235–243.

Sweet, A.L. and J.R. Wilson (1988), 'Pitfalls in simulation-based evaluation of forecast monitoring schemes', *International Journal of Forecasting*, 4, 573–579.

Tashman, L. and J.M. Kruk (1996), 'The use of protocols to select exponential smoothing procedures: A reconsideration of forecasting competitions', *International Journal of Forecasting*, 12, 235–253.

Taylor, J.W. (2003), 'Exponential smoothing with a damped multiplicative trend', *International Journal of Forecasting*, 19, 273–289.

Williams, D.W. and D. Miller (1999), 'Level-adjusted exponential smoothing for modeling planned discontinuities', *International Journal of Forecasting*, 15, 273–289.

Winters, P.R. (1960), 'Forecasting sales by exponentially weighted moving averages', *Management Science*, 6, 324–342.

Yar, M. and C. Chatfield (1990), 'Prediction intervals for the Holt-Winters forecasting procedure', *International Journal of Forecasting*, 6, 127–137.

ARIMA

de Alba, E. (1993), 'Constrained forecasting in autoregressive time series models: A Bayesian analysis', *International Journal of Forecasting*, 9, 95–108.

Ariño, M.A. and P.H. Franses (2000), 'Forecasting the levels of vector autoregressive log-transformed time series', *International Journal of Forecasting*, 16, 111–116.

Artis, M.J. and W. Zhang (1990), 'BVAR forecasts for the G-7', *International Journal of Forecasting*, 6, 349–362.

Ashley, R. (1988), 'On the relative worth of recent macroeconomic forecasts', *International Journal of Forecasting*, 4, 363–376.

Bianchi, L., J. Jarrett and T.C. Hanumara (1998), 'Improving forecasting for telemarketing centers by ARIMA modeling with interventions', *International Journal of Forecasting*, 14, 497–504.

Bidarkota, P.V. (1998), 'The comparative forecast performance of univariate and multivariate models: An application to real interest rate forecasting', *International Journal of Forecasting*, 14, 457–468.

Box, G.E.P. and G.M. Jenkins (1970), *Time Series Analysis Forecasting and Control*, San Francisco: Holden Day (revised ed. 1976).

Chevillon, G. and D.F. Hendry (2005), 'Non-parametric direct multi-step estimation for forecasting economic processes', *International Journal of Forecasting*, 21, 201–218.

Cholette, P.A. (1982), 'Prior information and ARIMA forecasting', *Journal of Forecasting*, 1, 375–383.

Cholette, P.A. and R. Lamy (1986), 'Multivariate ARIMA forecasting of irregular time series', *International Journal of Forecasting*, 2, 201–216.

Cummins, J.D. and G.L. Griepentrog (1985), 'Forecasting automobile insurance paid claims using econometric and ARIMA models', *International Journal of Forecasting*, 1, 203–215.

De Gooijer, J.G. and A. Klein (1991), 'On the cumulated multi-step-ahead predictions of vector autoregressive moving average processes', *International Journal of Forecasting*, 7, 501–513.

Dhrymes, P.J. and S.C. Peristiani (1988), 'A comparison of the forecasting performance of WEFA and ARIMA time series methods', *International Journal of Forecasting*, 4, 81–101.

Dhrymes, P.J. and D. Thomakos (1998), 'Structural VAR, MARMA and open economy models', *International Journal of Forecasting*, 14, 187–198.

Di Caprio, U., R. Genesio, S. Pozzi and A. Vicino (1983), 'Short term load forecasting in electric power systems: A comparison of ARMA models and extended Wiener filtering', *Journal of Forecasting*, 2, 59–76.

Downs, G.W. and D.M. Rocke (1983), 'Municipal budget forecasting with multivariate ARMA models', *Journal of Forecasting*, 2, 377–387.

Edlund, P-O. (1984), 'Identification of the multi-input Box-Jenkins transfer function model', *Journal of Forecasting*, 3, 297–308.

Edlund, P-O. and S. Karlsson (1993), 'Forecasting the Swedish unemployment rate. VAR vs. transfer function modelling', *International Journal of Forecasting*, 9, 61–76.

Engle, R.F. and C.W.J. Granger (1987), 'Co-integration and error correction: Representation, estimation, and testing', *Econometrica*, 55, 1057–1072.

Funke, M. (1990), 'Assessing the forecasting accuracy of monthly vector autoregressive models: The case of five OECD countries', *International Journal of Forecasting*, 6, 363–378.

Geriner, P.T. and J.K. Ord (1991), 'Automatic forecasting using explanatory variables: A comparative study', *International Journal of Forecasting*, 7, 127–140.

Geurts, M.D. and J.P. Kelly (1986), 'Forecasting retail sales using alternative models', *International Journal of Forecasting*, 2, 261–272.

Geurts, M.D. and J.P. Kelly (1990), 'Comments on: "In defense of ARIMA modeling" by D.J. Pack', *International Journal of Forecasting*, 6, 497–499.

Grambsch, P. and W.A. Stahel (1990), 'Forecasting demand for special telephone services: a case study', *International Journal of Forecasting*, 6, 53–64.

Guerrero, V.M. (1991), 'ARIMA forecasts with restrictions derived from a structural change', *International Journal of Forecasting*, 7, 339–347.

Gupta, S. (1987), Testing causality: Some caveats and a suggestion', *International Journal of Forecasting*, 3, 195–209.

Hafer, R.W. and R.G. Sheehan (1989), 'The sensitivity of VAR forecasts to alternative lag structures', *International Journal of Forecasting*, 5, 399–408.

Hansson, J., P. Jansson and M. Löf (2005), 'Business survey data: Do they help in forecasting GDP growth?", *International Journal of Forecasting*, 21, 377–389.

Harris, J.L. and L-M. Liu (1993), 'Dynamic structural analysis and forecasting of residential electricity consumption', *International Journal of Forecasting*, 9, 437–455.

Hein, S. and R.E. Spudeck (1988), 'Forecasting the daily federal funds rate', *International Journal of Forecasting*, 4, 581–591.

Heuts, R.M.J. and J.H.J.M. Bronckers (1988), 'Forecasting the Dutch heavy truck market: A multivariate approach', *International Journal of Forecasting*, 4, 57–59.

Hill, G. and R. Fildes (1984), 'The accuracy of extrapolation methods: An automatic Box-Jenkins package SIFT', *Journal of Forecasting*, 3, 319–323.

Hillmer, S.C., D.F. Larcker and D.A. Schroeder (1983), 'Forecasting accounting data: A multiple time-series analysis', *Journal of Forecasting*, 2, 389–404.

Holden, K. and A. Broomhead (1990), 'An examination of vector autoregressive forecasts for the U.K. economy', *International Journal of Forecasting*, 6, 11–23.

Hotta, L.K. (1993), 'The effect of additive outliers on the estimates from aggregated and disaggregated ARIMA models', *International Journal of Forecasting*, 9, 85–93.

Kang, I-B. (2003), 'Multi-period forecasting using different models for different horizons: an application to U.S. economic time series data', *International Journal of Forecasting*, 19, 387–400.

Kim, J.H. (2003), 'Forecasting autoregressive time series with bias-corrected parameter estimators', *International Journal of Forecasting*, 19, 493–502.

Kling, J.L. and D.A. Bessler (1985), 'A comparison of multivariate forecasting procedures for economic time series', *International Journal of Forecasting*, 1, 5–24.

Koreisha, S.G. (1983), 'Causal implications: The linkage between time series and econometric modelling', *Journal of Forecasting*, 2, 151–168.

Krishnamurthi, L., J. Narayan and S.P. Raj (1989), 'Intervention analysis using control series and exogenous variable in a transfer function model: A case study', *International Journal of Forecasting*, 5, 21–27.

Kunst, R. and K. Neusser (1986), 'A forecasting comparison of some VAR techniques', *International Journal of Forecasting*, 2, 447–456.

Landsman, W.R. and A. Damodaran (1989), 'A comparison of quarterly earnings per share forecast using James-Stein and unconditional least squares parameter estimators', *International Journal of Forecasting*, 5, 491–500.

Layton, A., L.V. Defris and B. Zehnwirth (1986), 'An international comparison of economic leading indicators of telecommunication traffic', *International Journal of Forecasting*, 2, 413–425.

Ledolter, J. (1989), 'The effect of additive outliers on the forecasts from ARIMA models', *International Journal of Forecasting*, 5, 231–240.

Leone, R.P. (1987), 'Forecasting the effect of an environmental change on market performance: An intervention time-series', *International Journal of Forecasting*, 3, 463–478.

LeSage, J.P. (1989), 'Incorporating regional wage relations in local forecasting models with a Bayesian prior', *International Journal of Forecasting*, 5, 37–47.

LeSage, J.P. and M. Magura (1991), 'Using interindustry input-output relations as a Bayesian prior in employment forecasting models', *International Journal of Forecasting*, 7, 231–238.

Libert, G. (1984), 'The M-competition with a fully automatic Box-Jenkins procedure', *Journal of Forecasting*, 3, 325–328.

Lin, W.T. (1989), 'Modeling and forecasting hospital patient movements: Univariate and multiple time series approaches', *International Journal of Forecasting*, 5, 195–208.

Litterman, R.B. (1986), 'Forecasting with Bayesian vector autoregressions – Five years of experience', *Journal of Business and Economic Statistics*, 4, 25–38.

Liu, L-M. and M-W. Lin (1991), 'Forecasting residential consumption of natural gas using monthly and quarterly time series', *International Journal of Forecasting*, 7, 3–16.

Liu, T-R., M.E. Gerlow and S.H. Irwin (1994), 'The performance of alternative VAR models in forecasting exchange rates', *International Journal of Forecasting*, 10, 419–433.

Lütkepohl, H. (1986), 'Comparison of predictors for temporally and contemporaneously aggregated time series', *International Journal of Forecasting*, 2, 461–475.

Makridakis, S., A. Andersen, R. Carbone, R. Fildes, M. Hibon, R. Lewandowski, J. Newton, E. Parzen and R. Winkler (1982), 'The accuracy of extrapolation (time series) methods: results of a forecasting competition', *Journal of Forecasting*, 1, 111–153.

Meade, N. (2000), 'A note on the Robust Trend and ARARMA methodologies used in the M3 Competition', *International Journal of Forecasting*, 16, 517–519.

Meade, N. (2002), 'A comparison of the accuracy of short term foreign exchange forecasting methods', *International Journal of Forecasting*, 18, 67–83.

Mélard, G. and J-M. Pasteels (2000), 'Automatic ARIMA modeling including interventions, using time series expert software', *International Journal of Forecasting*, 16, 497–508.

del Moral, M.J. and M.J. Valderrama (1997), 'A principal component approach to dynamic regression models', *International Journal of Forecasting*, 13, 237–244.

Newbold, P. (1983), 'ARIMA model building and the time series analysis approach to forecasting', *Journal of Forecasting*, 2, 23–35.

Newbold, P., C. Agiakloglou and J. Miller (1994), 'Adventures with ARIMA software', *International Jour-*

nal of Forecasting, 10, 573–581.

Öller, L-E. (1985), 'Macroeconomic forecasting with a vector ARIMA model', *International Journal of Forecasting*, 1, 143–150.

Pack, D.J. (1990), 'Rejoinder to: Comments on: "In defense of ARIMA modeling" by M.D. Geurts and J.P. Kelly', *International Journal of Forecasting*, 6, 501–502.

Parzen, E. (1982), 'ARARMA models for time series analysis and forecasting', *Journal of Forecasting*, 1, 67–82.

Pflaumer, P. (1992), 'Forecasting US population totals with the Box-Jenkins approach', *International Journal of Forecasting*, 8, 329–338.

Poskitt, D.S. (2003), 'On the specification of cointegrated autoregressive moving-average forecasting systems', *International Journal of Forecasting*, 19, 503–519.

Poulos, L., A. Kvanli and R. Pavur (1987), 'A comparison of the accuracy of the Box-Jenkins method with that of automated forecasting methods', *International Journal of Forecasting*, 3, 261–267.

du Preez, J. and S.F. Witt (2003), 'Univariate versus multivariate time series forecasting: An application to international tourism demand', *International Journal of Forecasting*, 19, 435–451.

Quenouille, M.H. (1957), *The Analysis of Multiple Time-Series*, London: Griffin. (2nd ed. 1968).

Ribeiro Ramos, F.F. (2003), 'Forecasts of market shares from VAR and BVAR models: A comparison of their accuracy', *International Journal of Forecasting*, 119, 95–110.

Reimers, H-E. (1997), 'Forecasting of seasonal cointegrated processes', *International Journal of Forecasting*, 13, 369–380.

Riise, T. and D. Tjøstheim (1984), 'Theory and practice of multivariate ARMA forecasting', *Journal of Forecasting*, 3, 309–317.

Shoesmith, G.L. (1992), 'Non-cointegration and causality: Implications for VAR modeling', *International Journal of Forecasting*, 8, 187–199.

Shoesmith, G.L. (1995), 'Multiple cointegrating vectors, error correction, and forecasting with Litterman's model', *International Journal of Forecasting*, 11, 557–567.

Simkins, S. (1995), 'Forecasting with vector autoregressive (VAR) models subject to business cycle restrictions', *International Journal of Forecasting*, 11, 569–583.

Spencer, D.E. (1993), 'Developing a Bayesian vector autoregressive forecasting model' , *International Journal of Forecasting*, 9, 407–421.

Tashman, L.J. and M.L. Leach (1991), 'Automatic forecasting software: a survey and evaluation', *International Journal of Forecasting*, 7, 209–230.

Tashman, L.J. (2000), 'Out-of sample tests of forecasting accuracy: a tutorial and review', *International Journal of Forecasting*, 16, 437–450.

Tegene, A. and F. Kuchler (1994), 'Evaluating forecasting models of farmland prices', *International Journal of Forecasting*, 10, 65–80.

Texter, P.A. and J.K. Ord (1989), 'Forecasting using automatic identification procedures: A comparative analysis', *International Journal of Forecasting*, 5, 209–215.

Villani, M. (2001), 'Bayesian prediction with cointegrated vector autoregression', *International Journal of Forecasting*, 17, 585–605.

Wang, Z. and D.A. Bessler (2004), 'Forecasting performance of multivariate time series models with a full and reduced rank: an empirical examination', *International Journal of Forecasting*, 20, 683–695.

Weller, B.R. (1989), 'National indicator series as quantitative predictors of small region monthly employment levels', *International Journal of Forecasting*, 5, 241–247.

Wieringa, J.E. and C. Horváth (2005), 'Computing level-impulse responses of log-specified VAR systems', *International Journal of Forecasting*, 21, 279–289.

Seasonality

Albertson, K. and J. Aylen (1996), 'Modelling the Great Lake freeze: Forecasting and seasonality in the market for ferrous scrap', *International Journal of Forecasting*, 12, 345–359.

Bunn, D.W and A.I. Vassilopoulos (1993), 'Using group seasonal indices in multi-item short-term forecasting', *International Journal of Forecasting*, 9, 517–526.

Bunn, D.W and A.I. Vassilopoulos (1999), 'Comparison of seasonal estimation methods in multi-item short-term forecasting', *International Journal of Forecasting*, 15, 431–443.

Chen, C. (1997), 'Robustness properties of some forecasting methods for seasonal time series: A Monte Carlo study', *International Journal of Forecasting*, 13, 269–280.

Clements, M.P. and D.F. Hendry (1997), 'An empirical study of seasonal unit roots in forecasting', *International Journal of Forecasting*, 13, 341–355.

Cleveland, R.B., W.S. Cleveland, J.E. McRae and I. Terpenning (1990), 'STL: A seasonal-trend decomposition procedure based on Loess' (with discussion), *Journal of Official Statistics*, 6, 3–73.

Dagum, E.B. (1982), 'Revisions of time varying seasonal filters', *Journal of Forecasting*, 1, 173–187.

Findley, D.F., K.C. Wills and B.C. Monsell (2004), 'Seasonal adjustment perspectives on "Damping seasonal factors: shrinkage estimators for the X-12-ARIMA program" ', *International Journal of Forecasting*, 20, 551–556.

Franses, P.H. and A.B. Koehler (1998), 'A model selection strategy for time series with increasing seasonal variation', *International Journal of Forecasting*, 14, 405–414.

Franses, P.H. and G. Romijn (1993), 'Periodic integration in quarterly UK macroeconomic variables', *International Journal of Forecasting*, 9, 467–476.

Franses, P.H. and D. van Dijk (2005), 'The forecasting performance of various models for seasonality and nonlinearity for quarterly industrial production', *International Journal of Forecasting*, 21, 87–102.

Gómez, V. and A. Maravall (2001), 'Seasonal adjustment and signal extraction in economic time series', Chapter 8 in *A course in time series analysis*, ed. D. Peña, G.C. Tiao and R.S. Tsay, John Wiley & Sons: New York.

Herwartz, H. (1997), 'Performance of periodic error correction models in forecasting consumption data', *International Journal of Forecasting*, 13, 421–431.

Huot, G., K. Chiu and J. Higginson (1986), 'Analysis of revisions in the seasonal adjustment of data using X-11-ARIMA model-based filters', *International Journal of Forecasting*, 2, 217–229.

Hylleberg, S. and A.R. Pagan (1997), 'Seasonal integration and the evolving seasonals model', *International Journal of Forecasting*, 13, 329–340.

Hyndman, R.J. (2004), 'The interaction between trend and seasonality', *International Journal of Forecasting*, 20, 561–563.

Koehler, A.B. (2004), 'Comments on damped seasonal factors and decisions by potential users', *International Journal of Forecasting*, 20, 565–566.

Kulendran, N. and M.L. King (1997), 'Forecasting international quarterly tourist flows using error-correction and time-series models', *International Journal of Forecasting*, 13, 319–327.

Ladiray, D. and B. Quenneville (2004), 'Implementation issues on shrinkage estimators for seasonal factors within the X-11 seasonal adjustment method', *International Journal of Forecasting*, 20, 557–560.

Miller, D.M. and D. Williams (2003), 'Shrinkage estimators of time series seasonal factors and their effect on forecasting accuracy', *International Journal of Forecasting*, 19, 669–684.

Miller, D.M. and D. Williams (2004), 'Damping seasonal factors: Shrinkage estimators for seasonal factors within the X-11 seasonal adjustment method' (with commentary), *International Journal of Forecasting*, 20, 529–550.

Noakes, D.J., A.I. McLeod and K.W. Hipel (1985), 'Forecasting monthly riverflow time series', *International Journal of Forecasting*, 1, 179–190.

Novales, A. and R.F. de Fruto (1997), 'Forecasting with time periodic models: A comparison with time invariant coefficient models', *International Journal of Forecasting*, 13, 393–405.

Ord, J.K. (2004), 'Shrinking: when and how?', *International Journal of Forecasting*, 20, 567–568.

Osborn, D. (1990), 'A survey of seasonality in UK macroeconomic variables', *International Journal of Forecasting*, 6, 327–336.

Paap, R., P.H. Franses and H. Hoek (1997), 'Mean shifts, unit roots and forecasting seasonal time series', *International Journal of Forecasting*, 13, 357–368.

Pfeffermann, D., M. Morry and P. Wong (1995), 'Estimation of the variances of X-11 ARIMA seasonally adjusted estimators for a multiplicative decomposition and heteroscedastic variances', *International Journal of Forecasting*, 11, 271–283.

Quenneville, B., D. Ladiray and B. Lefrançois (2003), 'A note on Musgrave asymmetrical trend-cycle filters', *International Journal of Forecasting*, 19, 727–734.

Simmons, L.F. (1990), 'Time-series decomposition using the sinusoidal model', *International Journal of Forecasting*, 6, 485–495.

Taylor, A.M.R. (1997), 'On the practical problems of computing seasonal unit root tests', *International Journal of Forecasting*, 13, 307–318.

Wells, J.M. (1997), 'Modelling seasonal patterns and long-run trends in U.S. time series', *International Journal of Forecasting*, 13, 407–420.

Withycombe, R. (1989), 'Forecasting with combined seasonal indices', *International Journal of Forecasting*, 5, 547–552.

State space and structural models and the Kalman filter

Coomes, P.A. (1992), 'A Kalman filter formulation for noisy regional job data', *International Journal of Forecasting*, 7, 473–481.

Harrison, P.J. and C.F. Stevens (1976), 'Bayesian forecasting', *Journal of the Royal Statistical Society (B)*, 38, 205–247.

Harvey, A.C. (1984), 'A unified view of statistical forecasting procedures' (with discussion), *Journal of Forecasting*, 3, 245–283.

Harvey, A.C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge: Cambridge University Press.

Harvey, A.C. and R.D. Snyder (1990), 'Structural time series models in inventory control', *International Journal of Forecasting*, 6, 187–198.

Kalman, R.E. (1960), 'A new approach to linear filtering and prediction problems', *Transaction of the ASME – Journal of Basic Engineering*, 82D, 35–45.

Mittnik, S. (1990), 'Macroeconomic forecasting experience with balanced state space models', *International Journal of Forecasting*, 6, 337–345.

Patterson, K.D. (1995), 'Forecasting the final vintage of real personal disposable income: A state space approach', *International Journal of Forecasting*, 11, 395–405.

Proietti, T. (2000), 'Comparing seasonal components for structural time series models', *International Journal of Forecasting*, 16, 247–260.

Ray, W.D. (1989), 'Rates of convergence to steady state for the linear growth version of a dynamic linear model (DLM)', *International Journal of Forecasting*, 5, 537–545.

Schweppe, F. (1965), 'Evaluation of likelihood functions for Gaussian signals', *IEEE Transactions on Information Theory*, 11(1), 61–70.

Vinod, H.D. and P. Basu (1995), 'Forecasting consumption, income and real interest rates from alternative state space models', *International Journal of Forecasting*, 11, 217–231.

Nonlinear

Astatkie, T., D.G. Watts and W.E. Watt (1997), 'Nested threshold autoregressive (NeTAR) models', *International Journal of Forecasting*, 13, 105–116.

Balkin, S.D. and J.K. Ord (2000), 'Automatic neural network modeling for univariate time series', *International Journal of Forecasting*, 16, 509–515.

Bradley, M.D. and D.W. Jansen (2004), 'Forecasting with a nonlinear dynamic model of stock returns and industrial production', *International Journal of Forecasting*, 20, 321–342.

Boero, G. and E. Marrocu (2004), 'The performance of SETAR models: A regime conditional evaluation of point, interval and density forecasts', *International Journal of Forecasting*, 20, 305–320.

Brockwell, P.J. and R.J. Hyndman (1992), 'On continuous-time threshold autoregression', *International Journal of Forecasting*, 8, 157–173.

Callen, J.F., C.C.Y. Kwan, P.C.Y Yip and Y. Yuan (1996), 'Neural network forecasting of quarterly accounting earnings', *International Journal of Forecasting*, 12, 475–482.

Cao, L and A.S. Soofi (1999), 'Nonlinear deterministic forecasting of daily dollar exchange rates', *International Journal of Forecasting*, 15, 421–430.

Cecen, A.A. and C. Erkal (1996), 'Distinguishing between stochastic and deterministic behavior in high frequency foreign rate returns: Can non-linear dynamics help forecasting', *International Journal of Forecasting*, 12, 465–473.

Chatfield, C. (1993), 'Neural network: forecasting breakthrough or passing fad?', *International Journal of Forecasting*, 9, 1–3.

Chatfield, C. (1995), 'Positive or negative', *International Journal of Forecasting*, 11, 501-502.

Church, K.B. and S.P. Curram (1996), 'Forecasting consumers' expenditure: A comparison between econometric and neural network models', *International Journal of Forecasting*, 12, 255–267.

Clements, M.P. and J. Smith (1997), 'The performance of alternative methods for SETAR models', *International Journal of Forecasting*, 13, 463–475.

Clements, M.P., P.H. Franses and N.R. Swanson (2004), 'Forecasting economic and financial time-series with non-linear models', *International Journal of Forecasting*, 20, 169–183.

Dahl, C.M. and S. Hylleberg (2004), 'Flexible regression models and relative forecast performance', *International Journal of Forecasting*, 20, 201–217.

Darbellay, G.A. and M. Slama (2000), 'Forecasting the short-term demand for electricity: Do neural networks stand a better chance?', *International Journal of Forecasting*, 16, 71–83.

De Gooijer, J.G. and V. Kumar (1992), 'Some recent developments in non-linear time series modelling, testing and forecasting', *International Journal of Forecasting*, 8, 135–156.

De Gooijer, J.G. and A. Vidiella-i-Anguera (2004), 'Forecasting threshold cointegrated systems', *International Journal of Forecasting*, 20, 237–253.

Enders, W. and B. Falk (1998), 'Threshold-autoregressive, median-unbiased, and cointegration tests of purchasing power parity', *International Journal of Forecasting*, 14, 171–186.

Fernández-Rodríguez, F., S. Sosvilla-Rivero and J. Andrada-Félix (1999), 'Exchange-rate forecasts with simultaneous nearest-neighbour methods; evidence from the EMS', *International Journal of Forecasting*, 15, 383–392.

Franses, P.H., R. Paap and B. Vroomen (2004), 'Forecasting unemployment using an autoregression with censored latent effects parameters', *International Journal of Forecasting*, 20, 255–271.

Ghiassi, M., H. Saidane and D.K. Zimbra (2005), 'A dynamic artificial neural network model for forecasting series events', *International Journal of Forecasting*, 21, 341–362.

Gorr, W. (1994), 'Research prospective on neural network forecasting', *International Journal of Forecasting*, 10, 1–4.

Gorr, W., D. Nagin and J. Szczypula (1994), 'Comparative study of artificial neural network and statistical models for predicting student grade point averages', *International Journal of Forecasting*, 10, 17–34.

Granger, C.W.J. and T. Teräsvirta (1993), *Modelling Nonlinear Economic Relationships*, Oxford: Oxford University Press.

Hamilton, J.D. (2001), 'A parametric approach to flexible nonlinear inference', *Econometrica*, 69, 537–573.

Heravi, S., D.R. Osborn and C.R. Birchenhall (2004), 'Linear versus neural network forecasting for European industrial production series', *International Journal of Forecasting*, 20, 435–446.

Herwartz, H. (2001), 'Investigating the JPY/DEM-rate: arbitrage opportunities and a case for asymmetry', *International Journal of Forecasting*, 17, 231–245.

Hill, T., L. Marquez, M. O'Connor and W. Remus (1994), 'Artificial neural network models for forecasting and decision making', *International Journal of Forecasting*, 10, 5–15.

Lisi, F. and A. Medio (1997), 'Is a random walk the best exchange rate predictor?' *International Journal of Forecasting*, 13, 255–267.

Ludlow, J. and W. Enders (2000), 'Estimating non-linear ARMA models using Fourier coefficients', *International Journal of Forecasting*, 16, 333–347.

Marcellino, M. (2004), 'Forecasting EMU macroeconomic variables', *International Journal of Forecasting*, 20, 359–372.

Olson, D. and C. Mossman (2003), 'Neural network forecasts of Canadian stock returns using accounting ratios', *International Journal of Forecasting*, 19, 453–465.

Poskitt, D.S. and A.R. Tremayne (1986), 'The selection and use of linear and bilinear time series models', *International Journal of Forecasting*, 2, 101–114.

Qi, M. (2001), 'Predicting US recessions with leading indicators via neural network models', *International Journal of Forecasting*, 17, 383–401.

Sarantis, N. (2001), 'Nonlinearities, cyclical behaviour and predictability in stock markets: International evidence', *International Journal of Forecasting*, 17, 459–482.

Swanson, N.R. and H. White (1997), 'Forecasting economic time series using flexible versus fixed specification and linear versus nonlinear econometric models', *International Journal of Forecasting*, 13, 439–461.

Tkacz, G. (2001), 'Neural network forecasting of Canadian GDP growth', *International Journal of Forecasting*, 17, 57–69.

Tong, H. (1978), 'On a threshold model'. In *Pattern recognition and signal processing* (ed. C.H. Chen). Amsterdam: Sijthoff and Noordhoff.

Volterra, V. (1930), *Theory of Functionals and of Integro-differential Equations*, Dover, New York.

Wiener, N. (1958), *Non-linear Problems in Random Theory*, London: Wiley.

Zhang, G., B.E. Patuwo and M.Y. Hu (1998), 'Forecasting with artificial networks: The state of the art', *International Journal of Forecasting*, 14, 35–62.

Long memory

Andersson, M.K. (2000), 'Do long-memory models have long memory?', *International Journal of Forecasting*, 16, 121–124.

Baillie, R.T. and S-K. Chung (2002), 'Modeling and forecasting from trend-stationary long memory models with applications to climatology', *International Journal of Forecasting*, 18, 215–226.

Beran, J., Y. Feng, S. Ghosh and P. Sibbertsen (2002), 'On robust local polynomial estimation with long-memory errors', *International Journal of Forecasting*, 18, 227–241.

Bhansali, R.J. and P.S. Kokoszka (2002), 'Computation of the forecast coefficients for multistep prediction of long-range dependent time series', *International Journal of Forecasting*, 18, 181–206.

Franses, P.H. and M. Ooms (1997), 'A periodic long-memory model for quarterly UK inflation', *International Journal of Forecasting*, 13, 117–126.

Granger, C.W.J. and R. Joyeux (1980), 'An introduction to long memory time series models and fractional differencing', *Journal of Time Series Analysis*, 1, 15–29.

Hurvich, C.M. (2002), 'Multistep forecasting of long memory series using fractional exponential models', *International Journal of Forecasting*, 18, 167–179.

Man, K.S. (2003), 'Long memory time series and short term forecasts', *International Journal of Forecasting*, 19, 477–491.

Öller, L-E. (1985), 'How far can changes in general business activity be forecasted?', *International Journal of Forecasting*, 1, 135–141.

Ravishanker, N. and B.K. Ray (2002), 'Bayesian prediction for vector ARFIMA processes', *International Journal of Forecasting*, 18, 207–214.

Ray, B.K. (1993), 'Long-range forecasting of IBM product revenues using a seasonal fractionally differenced ARMA model', *International Journal of Forecasting*, 9, 255–269.

Smith, J. and S. Yadav (1994), 'Forecasting costs incurred from unit differencing fractionally integrated processes', *International Journal of Forecasting*, 10, 507–514.

Souza, L.R. and J. Smith (2002), 'Bias in the memory for different sampling rates', *International Journal of Forecasting*, 18, 299–313.

Souza, L.R. and J. Smith (2004), 'Effects of temporal aggregation on estimates and forecasts of fractionally integrated processes: A Monte-Carlo study', *International Journal of Forecasting*, 20, 487–502.

ARCH/GARCH

Awartani, B.M.A. and V. Corradi (2005), 'Predicting the volatility of the S&P-500 stock index via GARCH models: the role of asymmetries', *International Journal of Forecasting*, 21, 167–183.

Bollerslev, T., R.Y. Chou and K.F. Kroner (1992), 'ARCH modeling in finance: a review of the theory and empirical evidence', *Journal of Econometrics*, 52, 5–59.

Brooks, C., S.P. Burke and G. Persand (2001), 'Benchmarks and the accuracy of GARCH model estimation', *International Journal of Forecasting*, 17, 45–56.

Engle, R.F. (1982), 'Autoregressive conditional heteroscedasticity with estimates of the variance of the United Kingdom inflation', *Econometrica*, 50, 987–1008.

Engle, R.F. and V. Ng (1993), 'Measuring and testing the impact of news on volatility', *Journal of Finance*, 48, 1749–1778.

Franses, P.H. and H. Ghijsels (1999), 'Additive outliers, GARCH and forecasting volatility', *International Journal of Forecasting*, 15, 1–9.

Galbraith, J.W. and T. Kisinbay (2005), 'Content horizons for conditional variance forecasts', *International Journal of Forecasting*, 21, 249–260.

Hentschel, L. (1995), 'All in the family: nesting symmetric and asymmetric GARCH models', *Journal of Financial Economics*, 39, 71–104.

Pagan, A. (1996), 'The econometrics of financial markets', *Journal of Empirical Finance*, 3, 15–102.

Sabbatini, M. and O. Linton (1998), 'A GARCH model of the implied volatility of the Swiss market index from option prices', *International Journal of Forecasting*, 14, 199–213.

Count data forecasting

Brännäs, K. (1995), 'Prediction and control for a time-series count data model', *International Journal of Forecasting*, 11, 263–270.

Brännäs, K., J. Hellström and J. Nordström (2002), 'A new approach to modelling and forecasting monthly guest nights in hotels', *International Journal of Forecasting*, 18, 19–30.

Croston, J.D. (1972), 'Forecasting and stock control for intermittent demands', *Operational Research Quarterly*, 23, 289–303.

Diebold, F.X., T.A. Gunther and A.S. Tay (1998), 'Evaluating density forecasts, with applications to financial risk management', *International Economic Review*, 39, 863–883.

Freeland, R.K. and B.P.M. McCabe (2004), 'Forecasting discrete valued low count time series', *International Journal of Forecasting*, 20, 427–434.

Grunwald, G.K., R.J. Hyndman, L.M. Tedesco and R.L. Tweedie (2000), 'Non-Gaussian conditional linear AR(1) models', *Australian and New Zealand Journal of Statistics*. 42, 479–495.

Johnston, F.R. and J.E. Boylan (1996), 'Forecasting intermittent demand: a comparative evaluation of Croston's method'. *International Journal of Forecasting*, 12, 297–298.

McCabe, B.P.M. and G.M. Martin (2005), 'Bayesian predictions of low count time series', *International Journal of Forecasting*, 21, 315–330.

Syntetos, A.A. and J.E. Boylan (2005), 'The accuracy of intermittent demand estimates', *International Journal of Forecasting*, 21, 303–314.

Willemain, T.R., C.N. Smart, J.H. Shockor and P.A. DeSautels (1994), 'Forecasting intermittent demand in manufacturing: a comparative evaluation of Croston's method', *International Journal of Forecasting*, 10, 529–538.

Willemain, T.R., C.N. Smart and H.F. Schwarz (2004), 'A new approach to forecasting intermittent demand for service parts inventories', *International Journal of Forecasting*, 20, 375–387.

Forecast evaluation and accuracy measures

Ahlburg, D.A., C. Chatfield, S.J. Taylor, P.A. Thompson, R.L. Winkler, A.H. Murphy, F. Collopy and R. Fildes (1992), 'A commentary on error measures', *International Journal of Forecasting*, 8, 99–111.

Armstrong, J.S. and F. Collopy (1992), 'Error measures for generalizing about forecasting methods: Empirical comparisons', *International Journal of Forecasting*, 8, 69–80.

Chatfield, C. (1988), 'Editorial: Apples, oranges and mean square error', *International Journal of Forecasting*, 4, 515–518.

Diebold, F.X. and R.S. Mariano (1995), 'Comparing Predictive Accuracy', *Journal of Business & Economic Statistics*, 13, 253–263.

Fildes, R. and S. Makridakis (1988), 'Forecasting and loss functions', *International Journal of Forecasting*, 4, 545–550.

Fildes, R. (1992), 'The evaluation of extrapolative forecasting methods', *International Journal of Forecasting*, 8, 81–98.

Fildes, R., M. Hibon, S. Makridakis and N. Meade (1998), 'Generalising about univariate forecasting methods: further empirical evidence', *International Journal of Forecasting*, 14, 339–358.

Flores, B. (1989), 'The utilization of the Wilcoxon test to compare forecasting methods: A note', *International Journal of Forecasting*, 5, 529–535.

Goodwin, P. and R. Lawton (1999), 'On the asymmetry of the symmetric MAPE', *International Journal of Forecasting*, 15, 405–408.

Granger, C.W.J. and Y. Jeon (2003a), 'A time-distance criterion for evaluating forecasting models', *International Journal of Forecasting*, 19, 199–215.

Granger, C.W.J. and Y. Jeon (2003b), 'Comparing forecasts of inflation using time distance', *International Journal of Forecasting*, 19, 339–349. (Corrigendum: Vol 19:4, p. 767).

Harvey, D., S. Leybourne and P. Newbold (1997), 'Testing the equality of prediction mean squared errors', *International Journal of Forecasting*, 13, 281–291.

Koehler, A.B. (2001), 'The asymmetry of the sAPE measure and other comments on the M3-Competition', *International Journal of Forecasting*, 17, 570–574.

Mahmoud, E. (1984), 'Accuracy in forecasting: A survey', *Journal of Forecasting*, 3, 139–159.

Makridakis, S., A. Andersen, R. Carbone, R. Fildes, M. Hibon, R. Lewandowski, J. Newton, E. Parzen and R. Winkler (1982), 'The accuracy of extrapolation (time series) methods: results of a forecasting competition', *Journal of Forecasting*, 1, 111–153.

Makridakis, S. (1993), 'Accuracy measures: theoretical and practical concerns', *International Journal of Forecasting*, 9, 527–529.

Makridakis, S. and M. Hibon (2000), 'The M3-Competition: results, conclusions and implications', *International Journal of Forecasting*, 16, 451–476.

Makridakis, S., S.C. Wheelwright and R.J. Hyndman (1998), *Forecasting: methods and applications*, John Wiley & Sons: New York.

McCracken, M.W. (2004), 'Parameter estimation and tests of equal forecast accuracy between non-nested models', *International Journal of Forecasting*, 20, 503–514.

Sullivan, R., A. Timmermann and H. White (2003), 'Forecast evaluation with shared data sets', *International Journal of Forecasting*, 19, 217–227.

Theil, H. (1966), *Applied Economic Forecasting*, North Holland: Amsterdam.

Thompson, P.A. (1990), 'An MSE statistic for comparing forecast accuracy across series', *International Journal of Forecasting*, 6, 219–227.

Thompson, P.A. (1991), 'Evaluation of the M-competition forecasts via log mean squared error ratio', *International Journal of Forecasting*, 7, 331–334.

Wun, L-M. and W.L. Pearn (1991), 'Assessing the statistical characteristics of the mean absolute error of forecasting', *International Journal of Forecasting*, 7, 335–337.

Combining

Aksu, C. and S. Gunter (1992), 'An empirical analysis of the accuracy of SA, OLS, ERLS and NRLS combination forecasts', *International Journal of Forecasting*, 8, 27–43.

Bates, J.M. and C.W.J. Granger (1969), 'Combination of forecasts', *Operations Research Quarterly*, 20, 451–468.

Bunn, D.W. (1985), 'Statistical efficiency in the linear combination of forecasts', *International Journal of Forecasting*, 1, 151–163.

Clemen, R.T. (1989), 'Combining forecasts: A review and annotated biography' (with discussion) *International Journal of Forecasting*, 5, 559–583.

Deutsch, M., C.W.J. Granger and T. Teräsvirta (1994), 'The combination of forecasts using changing weights', *International Journal of Forecasting*, 10, 47–57.

Diebold, F.X. and P. Pauly (1990), 'The use of prior information in forecast combination', *International Journal of Forecasting*, 6, 503–508.

Fang, Y. (2003), 'Forecasting combination and encompassing tests', *International Journal of Forecasting*, 19, 87–94.

Fiordaliso, A. (1998), 'A nonlinear forecast combination method based on Takagi-Sugeno fuzzy systems', *International Journal of Forecasting*, 14, 367–379.

Granger, C.W.J. and R. Ramanathan (1984), 'Improved methods of combining forecasts', *Journal of Forecasting*, 3, 197–204.

Gunter, S.I. (1992), 'Nonnegativity restricted least squares combinations', *International Journal of Forecasting*, 8, 45–59.

Hendry, D.F. and M.P. Clements (2002), 'Pooling of forecasts', *Econometrics Journal*, 5, 1–31.

Hibon, M. and T. Evgeniou (2005), 'To combine or not to combine: selecting among forecasts and their combinations', *International Journal of Forecasting*, 21, 15–24.

Kamstra, M. and P. Kennedy (1998), 'Combining qualitative forecasts using logit', *International Journal of Forecasting*, 14, 83–93.

de Menezes, L.M. and D.W. Bunn (1998), 'The persistence of specification problems in the distribution of combined forecast errors', *International Journal of Forecasting*, 14, 415–426.

Miller, S.M., R.T. Clemen and R.L. Winkler (1992), 'The effect of nonstationarity on combined forecasts', *International Journal of Forecasting*, 7, 515–529.

Newbold, P. and C.W.J. Granger (1974), 'Experience with forecasting univariate time series and the combination of forecasts', *Journal of the Royal Statistical Society (A)*, 137, 131–165.

Taylor, J.W. and D.W. Bunn (1999), 'Investigating improvements in the accuracy of prediction intervals for combinations of forecasts: A simulation study', *International Journal of Forecasting*, 15, 325–339.

Terui, N. and H.K. van Dijk (2002), 'Combined forecasts from linear and nonlinear time series models', *International Journal of Forecasting*, 18, 421–438.

Winkler, R.L. and S. Makridakis (1983), 'The combination of forecasts', *Journal of the Royal Statistical Society (A)*, 146, 150–157.

Zou, H. and Y. Yang (2004), 'Combining time series models for forecasting', *International Journal of Forecasting*, 20, 69–84.

Prediction intervals and densities

Chatfield, C. and A.B. Koehler (1991), 'On confusing lead time demand with h-period-ahead forecasts', *International Journal of Forecasting*, 7, 239–240.

Clements, M.P. and J. Smith (2002), 'Evaluating multivariate forecast densities: A comparison of two approaches', *International Journal of Forecasting*, 18, 397–407.

Clements, M.P. and N. Taylor (2001), 'Bootstrapping prediction intervals for autoregressive models', *International Journal of Forecasting*, 17, 247–267.

Grigoletto, M. (1998), 'Bootstrap prediction intervals for autoregressions: Some alternatives', *International Journal of Forecasting*, 14, 447–456.

Hyndman, R.J. (1995), 'Highest density forecast regions for non-linear and non-normal time series models', *Journal of Forecasting*, 14, 431–441.

Kim, J.A. (1999), 'Asymptotic and bootstrap prediction regions for vector autoregression', *International Journal of Forecasting*, 15, 393–403.

Kim, J.A. (2004), 'Bootstrap prediction intervals for autoregression using asymptotically mean-unbiased estimators', *International Journal of Forecasting*, 20, 85–97.

Koehler, A.B. (1990), 'An inappropriate prediction interval', *International Journal of Forecasting*, 6, 557–558.

Lam, J.-P. and M.R. Veall (2002), 'Bootstrap prediction intervals for single period regression forecasts', *International Journal of Forecasting*, 18, 125–130.

Lefrançois, P. (1989), 'Confidence intervals for non-stationary forecast errors: Some empirical results for the series in the M-competition', *International Journal of Forecasting*, 5, 553–557.

Makridakis, S. and M. Hibon (1987), 'Confidence intervals: an empirical investigation of the series in the M-competition' , *International Journal of Forecasting*, 3, 489–508.

Masarotto, G. (1990), 'Bootstrap prediction intervals for autoregressions', *International Journal of Forecasting*, 6, 229–239.

Pascual, L., J. Romo and E. Ruiz (2001), 'Effects of parameter estimation on prediction densities: A bootstrap approach', *International Journal of Forecasting*, 17, 83–103.

Pascual, L., J. Romo and E. Ruiz (2005), 'Bootstrap prediction intervals for power-transformed time series', *International Journal of Forecasting*, 21, 219–236.

Reeves, J.J. (2005), 'Bootstrap prediction intervals for ARCH models', *International Journal of Forecasting*, 21, 237–248.

Tay, A.S. and K.F. Wallis (2000), 'Density forecasting: a survey', *Journal of Forecasting*, 19, 235–254.

Taylor, J.W. and D.W. Bunn (1999), ‘Investigating improvements in the accuracy of prediction intervals for combinations of forecasts: A simulation study’, *International Journal of Forecasting*, 15, 325–339.

Wallis, K.F. (1999), ‘Asymmetric density forecasts of inflation and the Bank of England’s fan chart’, *National Institute Economic Review*, 167, 106–112.

Wallis, K.F. (2003), ‘Chi-squared tests of interval and density forecasts, and the Bank of England’s fan charts’, *International Journal of Forecasting*, 19, 165–175.

A look to the future

Armstrong, J.S. (2001), ’Suggestions for further research’, *www.forecastingprinciples.com/researchers.html*.

Chatfield, C. (1988), ‘The future of time-series forecasting’, *International Journal of Forecasting*, 4, 411–419.

Clements, M.P. (2003), ‘Editorial: Some possible directions for future research’, *International Journal of Forecasting*, 19, 1–3.

Cogger, K.C. (1988), ‘Proposals for research in time series forecasting’, *International Journal of Forecasting*, 4, 403–410.

Dawes, R., R. Fildes, M. Lawrence and J.K. Ord (1994), ‘The past and the future of forecasting research’, *International Journal of Forecasting*, 10, 151–159.

De Gooijer, J.G. (1990), ‘Editorial: The role of time series analysis in forecasting: A personal view’, *International Journal of Forecasting*, 6, 449–451.

De Gooijer, J.G. and A. Gannoun (2000), ‘Nonparametric conditional predictive regions for time series’, *Computational Statistics & Data Analysis*, 33, 259–275.

Durbin, J. and S.J. Koopman (2001), *Time Series Analysis by State Space Methods*, Oxford: Oxford University Press.

Engle, R.F and S. Manganelli (2004), ‘CAViaR: Conditional autoregressive value at risk by regression quantiles’, *Journal of Business & Economic Statistics*, 22, 367–381.

Harvey, A.C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge: Cambridge University Press.

Koenker, R.W. and G.W. Bassett (1978), ‘Regression quantiles’, *Econometrica*, 46, 33–50.

Ord, J.K. (1988), ‘Future developments in forecasting: The time series connexion’, *International Journal of Forecasting*, 4, 389–401.

Polonik, W. and Q. Yao (2000), ‘Conditional minimum volume predictive regions for stochastic processes’, *Journal of the American Statistical Association*, 95, 509–519.

Ramsay, J.O. and B.W. Silverman (1997), *Functional Data Analysis*, New York: Springer-Verlag. (2nd ed. 2005).

Stock, J.H. and M.W. Watson (1999), ‘A comparison of linear and nonlinear models for forecasting macroeconomic time series’. In R.F. Engle and H. White (eds.), *Cointegration, Causality and Forecasting*, pp. 1–44. Oxford: Oxford University Press.

West, M. and J. Harrison (1989), *Bayesian Forecasting and Dynamic Models*, New York: Springer Verlag (2nd ed. 1997).

Yao, Q. and H. Tong (1995), ‘On initial-condition and prediction in nonlinear stochastic systems’, *Bulletin International Statistical Institute*, IP10.3, 395–412.