ISSN 1440-771X

MONASH UNIVERSITY

AUSTRALIA

DEPARTMENT OF ECONOMETRICS AND BUSINESS STATISTICS

Using Evolutionary Spectra to Forecast Time Series

Elizabeth Ann Maharaj

Working Paper 4/2003

Using Evolutionary Spectra to Forecast Time Series

Elizabeth Ann Maharaj

Monash University

Australia

Abstract

In this paper, an adaptive smoothing forecasting approach based on evolutionary spectra as developed by Rao and Shapiro (1970) is applied to the 3003 time series of various types and lengths used in the M3-Competition (Makridakis and Hibon, 2000). Comparisons of out-of-sample forecasts are made with other methods used in the M3-Competition via the symmetric mean absolute percentage error (SMAPE). It will be seen that this method does appear to perform very well when applied specifically to yearly, quarterly and monthly macro time series and to yearly and monthly demographic time series used in the competition.

Keywords: Evolutionary Spectra, Adaptive Smoothing, M3-Competition

JEL Classification: C1

1 Introduction

Forecasting of business, economic and financial time series is carried out almost exclusively in the time domain. Various studies including those by Newbold and Granger (1974), Makridakis and Hibon (1979), the M-Competition of Makridakis et al. (1982), Meese and Gweke (1984), the M2-Competition of Makridakis et al. (1993), Fildes et al. (1998) and more recently the M3-Competition of Makridakis and Hibon (2000) have compared various time domain forecasting methods.

Frequency domain methods involve the analysis of the spectral density function. This function describes how the variation in a time series may be accounted for by cyclic components at different frequencies. While forecasting methods in the frequency domain have been developed mainly for special purpose forecasting in the areas of meteorology (DeMaria et al., 1991, Krstanovic and Singh 1993), and artificial intelligence (Park, et al. 1999), they have been almost completely absent in business, financial and economic applications. An exception is a paper by Rao and Shapiro (1970), in which an adaptive smoothing forecasting method based on evolutionary spectra was developed. Priestley (1965) first considered the evolutionary spectra approach to analysing non-stationary time series. The evolutionary spectrum of a time series is obtained by estimating successive spectra of overlapping portions of the time series. Since changes in the structure of a time series are mirrored by changes in the spectral characteristics of series, it makes sense to take into account these spectral characteristics when forecasting a time series. Many of the time domain methods are dependent on a series being stationary, or being easily transformed to a stationary series. Other time domain methods are applicable only to specific classes of linear or non-linear models. Since evolutionary spectra can be estimated for any type of series, stationary or non-stationary, linear or non-linear, forecasting based on it will therefore be quite generalized.

In this paper an empirical study is conducted in which comparisons are made between the frequency domain forecasting method of adaptive smoothing that is based on evolutionary spectra, and other conventional time domain forecasting methods. While Rao and Shapiro (1970) proposed this frequency domain forecasting method of forecasting, no extensive study to date has been carried out to compare this method to the more popularly used time domain forecasting methods. In Section 2, we briefly describe the concept of the evolutionary spectral density function and the approach used to smooth the time series adaptively and in Section 3, we report the results of the applications to the 3003 time series.

2 Evolutionary Spectra and the Adaptive Approach

The adaptive approach to exponential smoothing allows the smoothing constant to be modified as changes in the structure of the time series occurs. These changes can be tracked by evolutionary spectra of the time series. Because the structure of non-stationary series changes over time, estimating a conventional spectrum will not be appropriate. In order to take into account these structural changes over time, evolutionary spectra which are successive spectra of overlapping portions of the time series are estimated. Rao and Shapiro (1970) describe this as viewing the series through a moving time window of fixed length and they contend that the length of the window must be long enough so that fairly stable estimates are obtainable for a reasonable number of spectral components, but not too long so that any fundamental change that may occur in the time series is not lost in the smoothing.

Let $\{X_t, t = 1, 2, \ldots\}$ be a discrete parameter stochastic process. If $\{X_t\}$ is stationary then the spectral density function or spectrum is defined by

$$f_X(\omega) = \frac{1}{\pi} \sum_{k=-\infty}^{\infty} \gamma_k e^{i\omega k} \qquad (2.1)$$

where γ_k is the covariance function of X_t and ω, the frequency, is in the range $(0,\pi)$. Given a stationary time series $\{x_t, t = 1, 2, \ldots, T\}$, a smoothed estimator of the spectrum is

$$\hat{f}_X(\omega) = \frac{1}{\pi} \sum_{k=-m}^{m} \lambda_k c_k e^{i\omega k} \qquad (2.2)$$

where c_k is the covariance function of x_t and λ_k is a suitably chosen lag window and $m < T$ is called the truncation point. The choice of lag window is discussed in Priestley (1966). If X_t is non-stationary, the spectrum can be estimated within each overlapping window. It is assumed that the time series is stationary within each window of fixed length. For example, if a series $\{x_t, t = 1, 2, \ldots, T\}$ where $T = 60$ is considered and it is decided to take $nw = 15$ windows, then each window would contain $T - nw + 1 = 46$ observations and the spectrum for the points $x_1, x_2, \ldots, x_{46}$ is estimated; the spectrum for the points $x_2, x_3, \ldots, x_{47}$ is estimated and so on until the spectrum for $x_{15}, x_{16}, \ldots, x_{60}$ is estimated. For each window the spectrum is estimated at a set number of frequencies, which is usually half the number of data points. Suppose that with each window the spectrum is estimated at twenty three frequency points, then these spectra can be displayed in the following table.

Table 1 **Evolutionary Spectra**

Series	Frequencies Points				
	ω_1	ω_2	.	.	ω_{23}
$x_1, x_2, \dots, x_{46}$	$\hat{f}_1(\omega_1)$	$\hat{f}_1(\omega_2)$	.	.	$\hat{f}_1(\omega_{23})$
$x_2, x_3, \dots, x_{47}$	$\hat{f}_2(\omega_1)$	$\hat{f}_2(\omega_2)$	.	.	$\hat{f}_2(\omega_{23})$
. . . . . .	.	.	.	.	.
. . . . . .	.	.	.	.	.
$x_{15}, x_{16}, \dots, x_{60}$	$\hat{f}_{15}(\omega_1)$	$\hat{f}_{15}(\omega_2)$	.	.	$\hat{f}_{15}(\omega_{23})$

The natural logarithm of the spectral estimates is taken because they have equal variance. Changes in the structure of the time series will be manifested in the differences in successive spectra, that is, in the vectors $\left| ln\, \hat{f}_t(\omega_k) - ln\, \hat{f}_{t+1}(\omega_k) \right|$ for all *t*.

Following the procedure of Rao and Shapiro (1970), a moving average of three $ln\, \hat{f}_t(\omega_k)$ values is calculated within a particular frequency slot and the difference from the moving averages obtained as follows:

$$\delta_{tk} = \frac{1}{3}\left(ln\, \hat{f}_{t-2}(\omega_k) + ln\, \hat{f}_{t-1}(\omega_k) + ln\, \hat{f}_t(\omega_k)\right) - ln\, \hat{f}_t(\omega_k).$$

Then the maximum change in the spectrum smoothed over three periods is obtained, that is

$$\Delta_t = max_k \left|\delta_{tk}\right|$$

If the value of Δ_t is small compared to its standard deviation, it implies that no distinct change occurred in the structure of the time series and a low value of the smoothing constant is appropriate. When Δ_t is large compared to its standard deviation, the value of the smoothing constant should be driven upward towards one.

By deriving the asymptotic distribution of $\Delta_t = max\left|\delta_{tk}\right|$, Rao and Shapiro (1970) use the following method to determine the values of the smoothing constants α_t

Let

$$\beta_t = b + c(\Delta_t/\sigma)^2$$

where σ is the standard deviation of δ_{tk}, and b and c are determined from the conditions

$$b + r_1^2 c = 0.67 \text{ and } b + r_2^2 c = 0.095$$

so that

$$\alpha_t = Max(0.1, min((exp\,\beta_t - 1), 1)) \qquad (2.1)$$

will range between 0.1 and 0.95. r_1 represents the value of $\frac{\Delta_t}{\sigma}$ that will trigger a change in α_t to 0.95 and r_2 represents the value of $\frac{\Delta_t}{\sigma}$ below which α_t remains at 0.1. r_1 and r_2 are chosen according to the asymptotic distribution of Δ_t. Rao and Shapiro (1970) show that

$$P(\Delta_t < \chi) = exp(-nP(|\delta_{tk}| > \chi)) \qquad (2.2)$$

where n is the number of frequency points within each window. Now since $ln\,\hat{f}_t(\omega_k)$ is asymptotically normally distributed, δ_{tk} being a linear combination of the $ln\,\hat{f}_t(\omega_k)$ values is also asymptotically normally distributed and $\frac{\delta_{tk}^2}{\sigma^2}$ is asymptotically distributed as chi-square with one degree of freedom.

Hence to find a point χ such that $P(\Delta_t < \chi) = 0.99$, it can be seen from Equation (2.2) that

$$P(|\delta_{tk}| > \chi) = \frac{-ln(0.99)}{n} .$$

This implies that

$$P\left(\frac{\delta_{tk}^2}{\sigma^2} > \frac{\chi^2}{\sigma^2}\right) = \frac{-ln(0.99)}{n} .$$

Hence depending on the value of n, the chi-square percentile point can be obtained. r_1 selected as this percentile point $\frac{\chi}{\sigma}$ means that α_t is set to 0.95 when $P(\Delta_t < \chi) = 0.99$. r_2 is selected as the percentile point $\frac{\chi}{\sigma}$ below which α_t remains at 0.1 when $P(\Delta_t > \chi) = 0.99$ which implies from Equation (2.2) that

$$P(|\delta_{tk}| > \chi) = \frac{-ln(0.01)}{n}.$$

Hence r_2 is below $\frac{\chi}{\sigma}$ such that

$$P\left(\frac{\delta_{tk}^2}{\sigma^2} > \frac{\chi^2}{\sigma^2}\right) = \frac{-ln(0.01)}{n}.$$

Clearly, the number of frequency points selected must be such that

$$0 \le \frac{-ln(0.99)}{n} \le 1 \qquad \text{and} \qquad 0 \le \frac{-ln(0.01)}{n} \le 1$$

So once α_t is determined from Equation (2.1), the adaptive exponential smoothed forecast of x_{t+1} at time t would be obtained from

$$F_{t+1} = \alpha_t x_t + (1-\alpha_t)F_t$$

where $F_1 = x_1$ and $t \ge 1$. In what follows, the smoothing constants will be determined using the above method.

3 Application to the M3-Competition Data

Table 2 is a replication of a table for Mikridakis & Hibon (2000) showing the classification of the 3003 time series that were used in the M3 competition.

Table 2 The classification of the 3003 time series used in the M3-Competition

Time Interval between successive observations	Type of Time Series Data						
	Micro	Industry	Macro	Finance	Demographic	Other	Total
Yearly	146	102	83	58	245	11	645
Quarterly	204	83	336	76	57		756
Monthly	474	334	312	145	111	52	1428
Other	4			29		141	174
Total	828	519	731	308	413	204	3003

The evolutionary spectral adaptive smoothing method of forecasting was applied to all 3003 series and the out-of-sample forecasts obtained for forecast horizons 1-6 for yearly data, 1-8 for quarterly data, 1-18 for monthly data and 1-8 for other data. This is in keeping with what was done in Mikridakis & Hibon (2000) so that comparisons can be made. The symmetric mean absolute percentage error defined as

$$\text{SMAPE} = \sum \frac{|x_t - F_t|}{(x_t - F_t)/2} * 100$$

where x_t is the actual time series value and F_t is it's forecast, was used to analyse the performance of this method and was compared to the SMAPE's obtained for each of the methods used in the M3-Competition. Since 20 was the minimum series length for yearly data was and 24 for quarterly data was, in order to ensure that there was a sufficient number of observations and hence there was a sufficient number of frequency points within each window, four, six and eight windows were considered in turn for each series within which the spectra were estimated. For monthly and other-period data since 66 was the minimum series length, four to twenty windows in increments of two were considered. The Parzen lag window was used in each case to smooth the spectral estimates with each time window. Some of the time series with their forecasts are shown in Figure 1 from where it can be seen that the forecasts track the time series quite well especially for the monthly data.

For the all period time series, there was very little difference between the SMAPE's for the different number of windows considered. The results when six windows were considered are given in Tables 3 to 6.

Table 3 Average SMAPE: yearly data

Forecast Horizon	146 Micro	102 Industry	83 Macro	58 Finance	245 Demo	11 Other	Average
1	16.15	10.81	3.20	24.36	5.89	20.10	13.42
2	20.74	14.85	5.75	34.35	9.07	19.79	17.42
3	25.32	17.45	9.90	33.30	11.86	17.40	19.20
4	28.02	19.48	12.80	36.02	14.21	25.07	22.60
5	31.80	22.23	16.85	34.85	15.95	25.08	24.46
6	37.41	26.81	20.33	37.84	17.26	26.99	27.77

Table 4 Average SMAPE: quarterly[1] data

Forecast Horizon	204 Micro	83 Industry	336 Macro	76 Finance	57 Demo	Average
1	11.96	11.81	3.82	8.92	5.28	8.36
2	13.92	9.56	4.46	11.39	7.11	9.29
3	15.17	9.22	5.20	12.01	8.86	10.09
4	16.31	13.11	5.74	12.62	11.03	11.76
5	18.42	16.01	6.63	15.98	13.43	14.09
6	18.44	12.42	7.45	19.51	13.96	14.36
7	20.38	12.38	8.48	19.79	16.64	15.53
8	21.15	14.84	8.84	18.47	17.01	16.06

Table 5 Average SMAPE: monthly[2] data

Forecast Horizon	474 Micro	334 Industry	312 Macro	145 Finance	111 Demo	52 Other	Average
1	21.65	9.76	3.62	8.70	5.96	19.79	11.58
2	21.94	11.64	4.54	9.94	6.82	22.79	12.94
3	22.15	12.23	5.77	9.90	7.36	31.86	14.88
4	21.51	13.21	7.06	10.50	8.15	33.50	15.65
5	21.18	14.05	7.03	13.20	7.89	32.09	15.90
6	21.58	13.83	7.46	13.42	8.32	29.10	15.62
7	23.35	14.05	7.42	14.22	7.92	25.34	15.38
8	25.54	15.06	7.63	15.58	8.78	20.60	15.53
9	22.70	15.14	8.37	15.53	8.80	18.24	14.80
10	21.28	14.77	8.68	15.91	7.80	15.51	13.99
11	21.85	13.97	8.62	16.11	8.06	17.62	14.37
12	22.87	14.51	8.91	17.36	9.86	19.15	15.44
13	24.13	15.76	9.59	18.97	10.40	28.11	17.83
14	23.88	16.52	10.29	20.79	11.06	31.52	19.01
15	23.63	16.95	11.37	20.17	11.85	39.20	20.53
16	23.41	17.00	11.50	20.84	12.23	40.70	20.95
17	25.86	16.61	12.06	22.58	12.55	38.18	21.31
18	28.07	17.12	12.54	23.26	11.93	36.01	21.49

[1] For both quarterly and monthly data, forecasts where obtained by applying the method directly to the series.
[2] However when a sample of each type of series was first deseasonalized , forecasts obtained and forecasts then reseasonalized, and it was found that there was very little difference between the SMAPE's for the two methods.

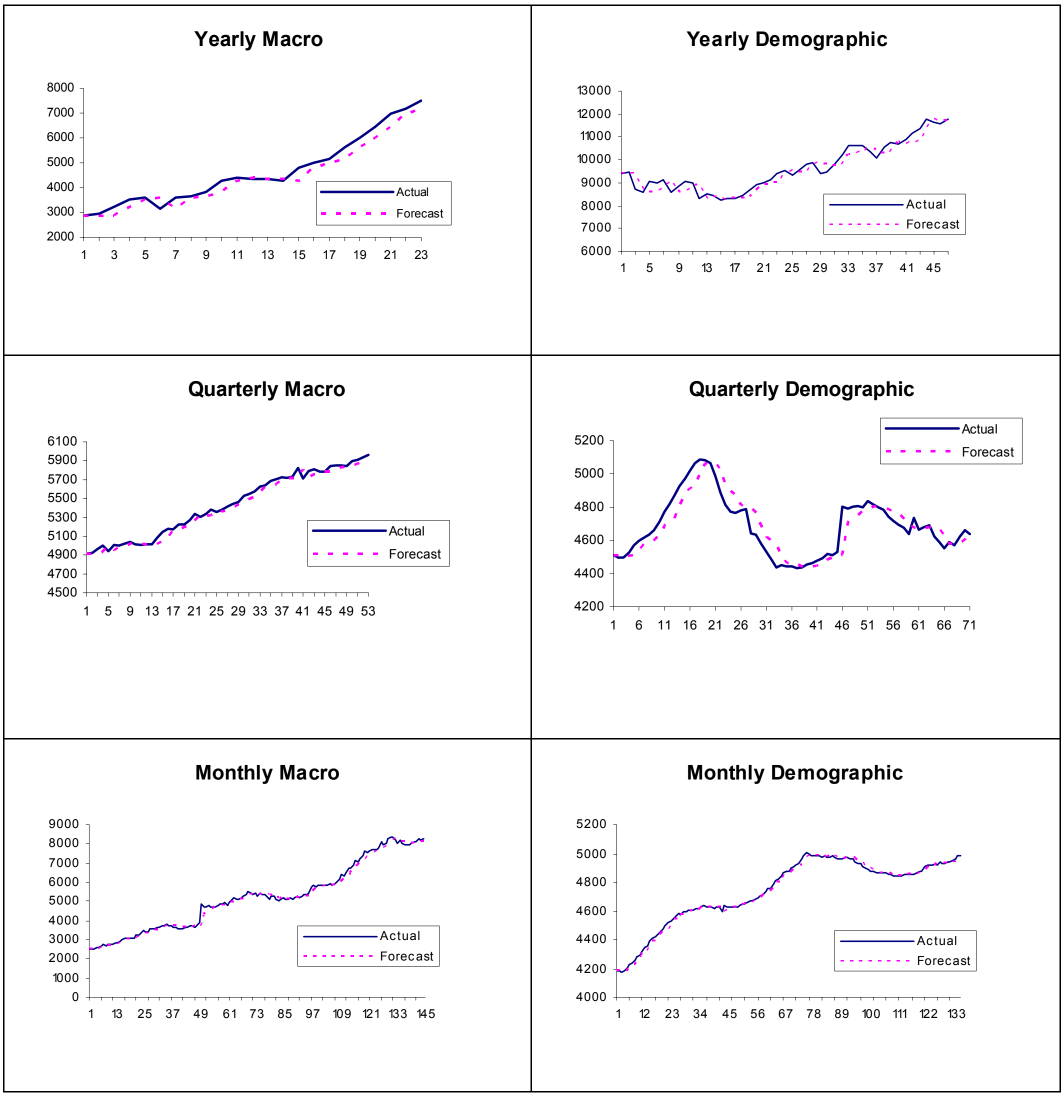
Table 6 Average SMAPE: other data

Forecast Horizon		4 Micro	29 Finance	141 Other	Average
1		8.78	1.12	2.76	4.22
2		12.98	1.60	3.58	6.06
3		14.89	1.88	3.97	6.92
4		17.53	2.23	5.33	8.36
5		19.84	2.53	5.89	9.42
6		21.23	2.89	8.14	10.75
7		24.29	3.09	9.14	12.17
8		29.13	3.42	10.84	14.46

This method appears to be doing quite well at forecasting macro time series and in that the SMAPE's for this type of series are consistently well below the average SMAPE's for all types for yearly, quarterly and monthly series for all of the methods considered in the M3-Competition[3] (see Tables 13 to 16 in Makridakis and Hibon, 2000). For yearly and monthly time intervals, the method is also doing quite well at forecasting demographic time series since again the SMAPE's for this type of series are consistently well below the average SMAPE's for all types for yearly and monthly series for all of the methods considered in the M3-Competition. Furthermore for monthly data, the average SMAPEs for this method for all types of series are comparable to those for the different methods used in the M3-Competition. For other-period financial data, this method performs very well since its SMAPE's at the various forecast horizons is lower than the average SMAPE's for other period data when all methods were used in the M3-Competition.

[3] Comparisons are made with the SMAPE's averaged across the various types of series from the M3-Competition because those are the only results made available by Makridakis and Hibon (2000).

Figure 1 Yearly, Quarterly and Monthly Macro and Demographic Time Series and Forecasts



4 Concluding Remarks

On average, for all types of yearly and quarterly time series , the evolutionary spectral adaptive smoothing method of forecasting does not perform as well as the other methods used in the M3-Competition. However when applied to monthly time series, its performance on average is comparable to that of other methods. The method does appear to perform very well when applied specifically to all-period macro time series and to yearly and monthly demographic time series considered in the competition. A possible explanation for this would be that generally the evolutionary spectra are better able to track the changes in structure of these types of series than the other type of series considered.

Acknowledgements

This study was supported in part by a research grant from the Faculty of Business and Economics and in part by funding from the Business & Economic Forecasting Unit of the Department of Econometrics and Business Statistics.

References

DeMaria, M., Aberson, S.D., Lord, S.J., (1992), A nested spectral model for hurricane tracking forecasting, *Monthly Weather Review*, 120, 1628.

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

Krstanovic, P.F. and Singh, V.P., (1993), A real-time flood forecasting model based on maximum-entropy spectral analysis, *Water Resources Management*, 7, 109-129.

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

Makridakis, S., Chatfiled, C., Hibon, M., Lawrence, M., Mills, T., Ord, K., Simmons, L.F., (1993), The M2-competition: a real-time judgmentally based forecasting study, *International Journal of Forecasting*, 9, 5-22.

Makridakis, S. and Hibon, M., (1979), Accuracy of forecasting; an empirical investigation (with discussion), *Journal of the Royal Statistical Society*, (A),142, 97-145.

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

Meese, R. and Gweke, J., (1984), A comparison of autoregressive univariate Forecasting procedures for macroeconomic time series, *Journal of Business and Economic Statistics*, 2, 191-200.

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

Park, D., Rilett, L.R. and Han, G., Spectral basis neural networks for real-time travel time forecasting (1999), *Journal of Transportation Engineering*, 125, 515.

Priestley, M.B., (1965), Evolutionary spectra and non-stationary processes, *Journal of the Royal Statistical Society*, (B), 27, 204-237.

Priestley, M.B. (1966), Design relations for non-stationary processes. *Journal of the Royal Statistical Society*, (B), 28, 228-240.

Rao, A.G. and Shapiro, A., (1970), Adaptive smoothing evolutionary spectra, *Management Science*, 17, 208-281.