Forecasting telecommunications data with linear models

Madden, Gary G and Tan, Joachim

Curtin University of Technology, School of Economics and Finance, Perth WA 6845, Australia, Curtin University of Technology, School of Economics and Finance, Perth WA 6845, Australia

2007

Forecasting Telecommunications Data with Linear Models*

Abstract

For telecommunication companies to successfully manage their business, companies rely on mapping future trends and usage patterns. However, the evolution of telecommunications technology and systems in the provision of services renders imperfections in telecommunications data and impinges on a company's' ability to properly evaluate and plan their business. ITU *Recommendation E.507* provides a selection of econometric models for forecasting these trends. However, no specific guidance is given. This paper evaluates whether simple extrapolation techniques in *Recommendation E.507* can generate accurate forecasts. Standard forecast error statistics—mean absolute percentage error, median absolute percentage error and percentage better—show the ARIMA, Holt and Holt-D models provide better forecasts than a random walk and other linear extrapolation methods.

Keywords: linear models, ITU Recommendations, telecommunications forecasting

Gary Madden
School of Economics and Finance, Curtin University of Technology
GPO Box U1987, Perth, WA 6845
Australia
Tel: 61-8-9266-4258
gary.madden@cbs.curtin.edu.au

Joachim Tan
School of Economics and Finance, Curtin University of Technology
Perth, WA 6845
Australia

This Draft: 2nd May 2006

* This is a preliminary draft—comments are welcome. Corresponding author is Gary Madden. Any errors and omissions are our own responsibility.

I. Introduction

In the current environment telecommunications industry boundaries are rapidly changing, markets are increasingly competitive, and company-based data are more fragmented and proprietary. Within this context, telecommunications companies must develop strategy to deal with a rapidly changing and expanding mix of telecommunications services, especially the emergence of substitutes and complements. That is, to successfully manage their business carriers must rely on data to monitor, analyse and optimize their systems to map future trends and use patterns. Forecasting is an integral input into such network traffic management, infrastructure optimization and planning, and the scenario planning process. Data-based forecasts, excluding the analysis, are constructed from data that is processed via statistical models from which inferences are drawn. The analysis of Grubesic and Murray (2005) addresses concerns related to the quality of available (spatial) data. In particular, Grubesic and Murray argue that, because of the complexity of telecommunications systems, analysts need to concern themselves with the impact of using imperfect spatial information. Finally, Grubesic and Murray develop a framework to address the sensitivity of spatial analyses to imperfect spatial data. By contrast, this study focuses on identifying appropriate statistical method given no information about the available data. This approach is in the spirit of the International Telecommunication Union (ITU) *Recommendation E.507*. The *Recommendation* provides an overview of existing mathematical techniques for forecasting that includes curve-fitting models, autoregressive models, autoregressive integrated moving average (ARIMA) models, state space models with Kalman filtering and regression models.[1] However, *Recommendation E.507* provides no guidance as to which models are the most appropriate for forecasting telecommunications series. This study contributes by examining

[1] *Recommendation E.507* also describes methods to evaluate and select an appropriate technique, depending on available data and forecast period.

the forecast performance of these extrapolation models by comparing forecasts against those from a random walk model.[2] In particular, the paper follows the format of a competition between simple models that do not require detailed domain knowledge. Namely, the comparison of forecast accuracy is based on telecommunications data via series pattern recognition. The paper is structured as follows. Section II describes the M3-competition telecommunications data. Section III reviews linear univariate forecast methods employed in this study, and discusses error statistics used to evaluate forecast performance. Section IV presents the forecast results and Section V concludes.

II. Data

Data used is acquired from the Institute of Forecasters Web site located at http://www.forecasters.org. M3-competition data are comprised of 3003 series. 149 series are telecommunications industry data, of which 29 are monthly series and 120 series are of unknown periodicity (and labelled 'other'). Monthly and 'other' series consist of 53 and 63 observations, respectively. All observations have strictly positive values. No additional information is provided by the Institute. A representative specimen of the monthly and 'other' data are shown in Fig. I and Fig. II.

<Insert Fig. I & Fig. II>

Grambsch and Stahel (1990) and Fildes (1992) find telecommunications data exhibit both non-stationary and strong negative trends.[3] Similarly, most of the monthly and 'other' series

[2] Standard univariate linear models are useful for this application as they are simple to implement and easily understood, and therefore used for commercial applications.
[3] These observations lead Grambsch and Stahel (1990) and Fildes (1992) to argue the simple exponential smoothing model is inappropriate for forecasting these data.

in this study exhibit negative trends. For the monthly series, only 1 of 29 series shows a positive trend. For the 'other' series, 115 series exhibits a negative trend, while only 5 series exhibit a positive trend.

<Insert Table I & Table II>

Average summary statistics for the monthly and 'other' series are presented in Table I and Table II, respectively. The tables indicate the mean, standard deviation, degree of skewness and number of outliers for the 'other' series is higher than those of the monthly series. The sample average of standard deviation for the monthly and 'other' series are 386.8 and 921.5, respectively. This indicates the spread of the 'other' series is about 2.4 times larger than that of the monthly series. The corresponding coefficients of variation (0.08 and 0.13) for both series also indicate the differences in the spread of the series.

Following Fildes (1992), frequency of outliers, strength of trend and degree of randomness for the data are analysed. Results are shown in Fig. V and Fig. VI.[4]

<Insert Fig. III and Fig. IV>

---

[4] An observation is defined as being an outlier when either $X_t < L_x - 1.5(U_x - L)$ or $X_t > L_x + 1.5(U_x - L_x)$ where L_x denotes the lower quartile and U_x is the upper quartile. The strength of trend is measured by the correlation between series (with outliers removed) and a time trend, with the absolute value of the trend indicating its strength. Randomness is measured by estimating the regression:

$$X'_t = \alpha + \beta t + \delta_1 X'_{t-1} + \delta_2 X'_{t-2} + \delta_3 X'_{t-3}$$

Where X'_t denotes the series X_t with outliers removed. $\overline{R}^2$ measures the variation explained by the model. A higher $\overline{R}^2$ indicates little randomness, while a relatively low $\overline{R}^2$ suggests a higher degree of randomness.

Figure III reveals 26 of the 29 monthly series contain only a single outlier, with 3 series containing more than 1 outlier. The ‘other’ series exhibit some similar properties. That is; Fig. IV shows that 116 of the 120 ‘other’ series contain only 1 outlier, 2 series contain 11 outliers and 2 series contains 14 outliers. The characteristics of the M3 telecommunications data appear to be homogenous and have similar properties to the telecommunications data analysed by Fildes (1992).

<Insert Fig. V & Fig. VI>

Figure V and Fig. VI show the strength of the trend for the M3 data. The monthly series have the stronger negative correlation with time. The histograms contained in Fig. V and Fig. VI also reveal a higher correlation for the monthly series.

<Insert Fig. VII & Fig. VIII>

Figure VII and Fig. VIII show both the monthly and ‘other’ series exhibit a low degree of randomness and have a strong positive correlation with a linear trend. The results suggests employing linear models to forecast the M3 telecommunications data is appropriate.

III. Forecast Models Applied

The univariate linear extrapolative techniques applied for forecasting are ARARMA (Parzen 1982), ARIMA, Holt, Holt-D, Holt-Winters, simple exponential smoothing (SES) and the robust trend (RT; Grambsch and Stahel 1990) models.[5] These forecast methods are employed as the models are proposed in *Recommendation E.507* and are shown to be reliable by Makridakis et al. (1993), Fildes et al. (1998), and Makridakis and Hibon (2000) by performing consistently in the M-competition studies. All models are estimated beginning at observation 6. This means parameter estimates for all models are estimated from observation 6 to observation 53, and observation 6 to observation 63 for the monthly and the 'other' series, respectively. To estimate the parameters of the ARARMA and ARIMA models, an automatic procedure with a maximum 5-period lag is employed. The automatic procedure estimates the parameters of all possible combinations of ARARMA and ARIMA models within the imposed lag limit generating 60 possible models for each ARARMA and ARIMA model, respectively.[6] A grid search is then performed on the generated ARARMA and ARIMA models to determine the optimal lag length. This is done by comparing the Akaike Information Criterion (AIC) statistic values of the generated models.[7] The 'best' ARARMA and ARIMA models are those that generate the lowest AIC. Holt, Holt-D, Holt-Winters and RT models have their lag lengths fixed and so do not require a grid search to select best model.[8] Following Makridakis and Hibon (2000), a best model—for both series and method—is used to forecast a maximum 18 observations and 8 observations ahead for the first forecast in the sequence, respectively, for the monthly and 'other' series. For the second forecast in the sequence, the data series expands by one period (one-step ahead) and

[5] Only the no trend, no seasonal version of the SES model is included in the analysis, i.e., the SES model used is $y_t = y_{t-1} + \alpha e_t$, with $\alpha = 0.3$.

[6] The maximum lags for the ARARMA model is 5 periods. The ARARMA is estimated by applying an AR model 1- and 2-period lags. Residuals are estimated with another ARMA model with lags of 1- to 3-periods to generate the ARARMA model.

[7] $$AIC = \log(\hat{\sigma}^2_{(p,q)}) + \frac{2m}{T}$$

where $\hat{\sigma}^2_{(p,q)}$ is the variance of the residuals of the estimated model. T is the number of observations and m is the number of parameters of the univariate model tested.

[8] Only the linear, no trend and non-seasonal versions of the Holt and Holt-W methods are considered. The Holt-D model is the exponentially smoothed version of the Holt model. Parameters for these models are estimated from data rather than being fixed arbitrarily.

forecasts are made for 17 observations and 7 observations ahead, respectively. Forecast accuracy measures employed are guided by Armstrong and Collopy (1992), viz., mean absolute percentage error (MAPE), median absolute percentage error (MdAPE) and percent better (PB). The error statistics are defined in the Appendix.[9]

IV. Forecasts

To identify the most accurate forecast model, aggregate results by method are compared using an out of sample forecast horizon for the monthly and ‘other’ series data. Following Makridakis and Hibon (2000), a maximum of 18 and 8 steps ahead are generated for monthly and ‘other’ data, respectively. Forecast accuracy is compared using MAPE, MdAPE and PB error statistics. Table III and Table IV present the MAPE and MdAPE results for the out-of-sample monthly forecasts. MAPE statistic show the Holt model as best forecast method for medium-and long-horizons (short-, intermediate- and long-horizons are considered), as it consistently the yields lowest percentage error when compared to sample data. The MdAPE statistic indicates the Holt-D model is the best method for medium- and long-horizons. For short-horizon forecasting (1-period ahead), the MAPE and MdAPE statistics suggest the ARIMA model is best model.

<Insert Table III, Table IV & Table V>

The results for the PB statistic are reported for monthly series in Table V. Similar to the MdAPE results of Table IV, Table V shows the ARIMA and the Holt-D model are best. Table V shows the ARIMA model is best at forecasting short- and long-horizons for 1-

[9] Mean square error measures are not used as they are scale dependent and sensitive to outliers.

period, 12-periods and 18-periods ahead, while the Holt-D model is best at forecasting intermediate- to long-horizons of 6-periods and 12-periods ahead.

Figure IX through Fig. XI illustrate forecasts for 1-period to 18-periods ahead for the best forecast methods using monthly data. The MAPE statistic in Fig. IX shows the Holt and RW models is best for most forecast horizons. At shorter horizons of 8-periods ahead or less, the MAPE statistic shows the ARIMA and Holt models generate the best forecasts.

The MdAPE statistic in Fig. X indicates the Holt model and Holt-W model typically forecast best for all horizons. The only exception is the ARIMA model forecast better than the Holt-W model at 12-periods ahead. The PB statistic in Fig. XI, show the Holt-D model and ARIMA model are the best forecast models for most forecast horizons, with the Holt-W model is equally accurate in forecasting horizons 6-periods and 9-periods ahead. That is, the MAPE, MdAPE and PB statistics suggest the Holt and Holt-D models provide an improvement over other linear extrapolation techniques and a random walk model in providing forecasts for monthly data. This indicates the Holt and Holt-D models are useful for establishing relatively accurate judgement-free projections of telecommunications data up to 18-periods ahead for monthly data. Holt and Holt-D models are superior for most forecast horizons. For short-horizon forecasts, the ARIMA model is best.

<Insert Fig. IX, Fig. X & Fig. XI>

Table VI and Table VII present the MAPE and MdAPE results of the out-of-sample forecasts for the ‘other’ series. The results in Table VI and Table VII are similar to the forecast error results of the monthly series. The MAPE statistic in Table VI shows the Holt model is best forecast method for medium- and long-horizons, and yields lowest percentage error when forecasting 4-periods, 6-periods and 8-periods ahead. For short horizons of 1-period ahead, the MAPE statistic shows the ARIMA and the Holt-D are the best for forecasting 'other' series.

The MdAPE error statistic presented in Table VII shows a similar result to the MAPE statistic. The Holt model is the best for forecasting 4-periods, 6-periods and 8-periods ahead. The MdAPE statistic also shows the Holt-D model is best for forecasting short- and intermediate-horizon periods, viz., 1-period and 4-periods ahead. An exception contained in Table VII shows the Holt-W model performs as well as the Holt and Holt-D models when forecasting 4-periods ahead.

<Insert Table VI, Table VII & Table VIII>

The PB statistic presented in Table VIII shows the Holt-D model is best for forecasting short- and intermediate-horizons of 1-period and 4-periods ahead for the 'other' series, while the Holt model is best suited for forecasting the intermediate to long-horizons of 6-periods and 8-periods ahead.

<Insert Fig. XII, Fig. XIII & Fig. XIV>

When forecasting the 'other' series, the MAPE and MdAPE statistics contained in Fig. XII and Fig. XIII shows the Holt, Holt-D and Holt-W methods provide an improvement over other linear extrapolation techniques and random walk model. The PB statistic in Fig. XIV show the Holt model and Holt-D models provide better forecasts for the 'other' series. The results suggest the Holt model, Holt-D model and Holt-W models are useful for establishing relatively accurate judgement-free projections of telecommunications data of up to 8-periods ahead for telecommunications data of unknown periodicity. The results of the Holt and Holt-D models are superior for most horizons.

<Insert Table IX, Table X & Table XI>

A summary of the results by error statistic and forecast horizon are presented in Table IX, Table X and Table XI. The tables indicate the best model for forecasting telecommunication series without any domain knowledge for the forecast horizons tested are the ARIMA, Holt and the Holt-D models.

V. Conclusion

This analysis intends to identify those of the linear models proposed in ITU's *Recommendation E.507* (to aid telecommunication companies in forecasting) provide the better forecasts when little domain information is available. In particular, the analysis covers situation whereby little or no information is available about the reliability or quality of data. Forecasts from the ARARMA, ARIMA and several SES model specifications are compared against those for a random walk model. The results show despite having no knowledge of the origin of the telecommunications series, the models can accurately forecast telecommunications data. This suggests imperfections that may be inherent within telecommunications data might not unduly affect the accuracy of generated telecommunication forecasts. Telecommunications data are often not well understood and are short in length. Hence, the use of these univariate models without assumptions *a priori* may be most appropriate. The study shows the linear extrapolation models perform well when employed to forecast monthly and 'other' M3 telecommunications series of unknown periodicity. The sample forecast accuracy based on MAPE and MdAPE statistics, and PB measure show the Holt and Holt-D models provides the most reliable forecasts without any knowledge of the underlying dynamics of the telecommunications data series, an outcome similar to the recommendations of the M3-competition. The results suggest further gains in forecast accuracy may not be worthwhile in employing more sophisticated nonlinear models, as the forecast performance of the univariate linear models applied can generate accurate forecasts than the random walk for intermediate and long-horizons future periods.

References

Armstrong, J. and Collopy, F. (1992), 'Error Measures for Generalizing about Forecast Methods: Empirical Comparisons', *International Journal of Forecasting*, 8, 69-80.

Fildes, R. (1992), 'The Evaluation of Extrapolative Forecasting Methods', *International Journal of Forecasting*, 8, 81-98.

Fildes, R., Hibon, M., Makridakis, S. and Meade, N. (1998), 'Generalising about Univariate Forecasting Methods: Further Empirical Evidence', *International Journal of Forecasting*, 14, 339-58.

Gardner, E. and McKenzie, E. (1985), 'Forecasting Trends in Time Series', *Management Science*, 31, 1237-1246.

Grambsch, P. and Stahel, W. (1990), 'Forecasting Demand for Special Services', *International Journal of Forecasting*, 6, 53-64.

Grubesic, T. and Murray, A. (2005), 'Geographies of Imperfection in Telecommunication Analysis', *Telecommunications Policy*, 29, 69-94.

Ionescu-Graff, A. (1982), 'A Sequential Projection Algorithm for Special Services Demand', *Bell Systems Technical Journal*, 61, 34-66.

Makridakis, S., Chatfield, C., Hibon, M., Lawrence, M., Mills, T., Ord, K. and Simmons, L. (1993), 'The M-2 Competition: A Real-time Judgmentally Based Forecasting Study', *International Journal of Forecasting*, 9, 5-23.

Makridakis, S. and Hibon, M. (2000), 'The M3-competition: Results, Conclusions and Implications', *International Journal of Forecasting*, 16, 451-76.

Pack, C. and Whitaker, B. (1982), 'Kalman Filter Models for Network Forecasting', *Bell System Technical Journal*, 61, 1-14.

Pankratz, A. (1983), *Forecasting with Univariate Box-Jenkins Models: Concepts and Cases*, John Wiley: New York.

Parzen, E. (1982), 'ARARMA Models for Time Series Analysis and Forecasting', *Journal of Forecasting*, 1, 67-82.

Appendix

To estimate the MAPE for method i and horizon h for series j, the absolute percentage error $APE_{i,h,j}$ is first calculated;

$$APE_{i,h,j} = \left| \frac{F_{i,h,j} - A_{h,j}}{A_{h,j}} \right|,$$

where $F_{i,h,j}$ is the forecast for method i for horizon h using series j and $A_{h,j}$ is the actual value at horizon h for series *j*. The $MAPE_{i,h,j}$ is then calculated by finding the mean of the

$APE_{i,h,j}$ series;

$$MAPE_{i,h,j} = mean\ (APE_{i,h,j}),$$

As none of the series observations are negative, the MAPE is appropriate. An advantage of applying the MAPE is that it is scale-invariant. To estimate the MdAPE error statistic for method i for series j from the $APE_{i,h,j}$ error statistic, the median of the $APE_{i,h,j}$ series is calculated:

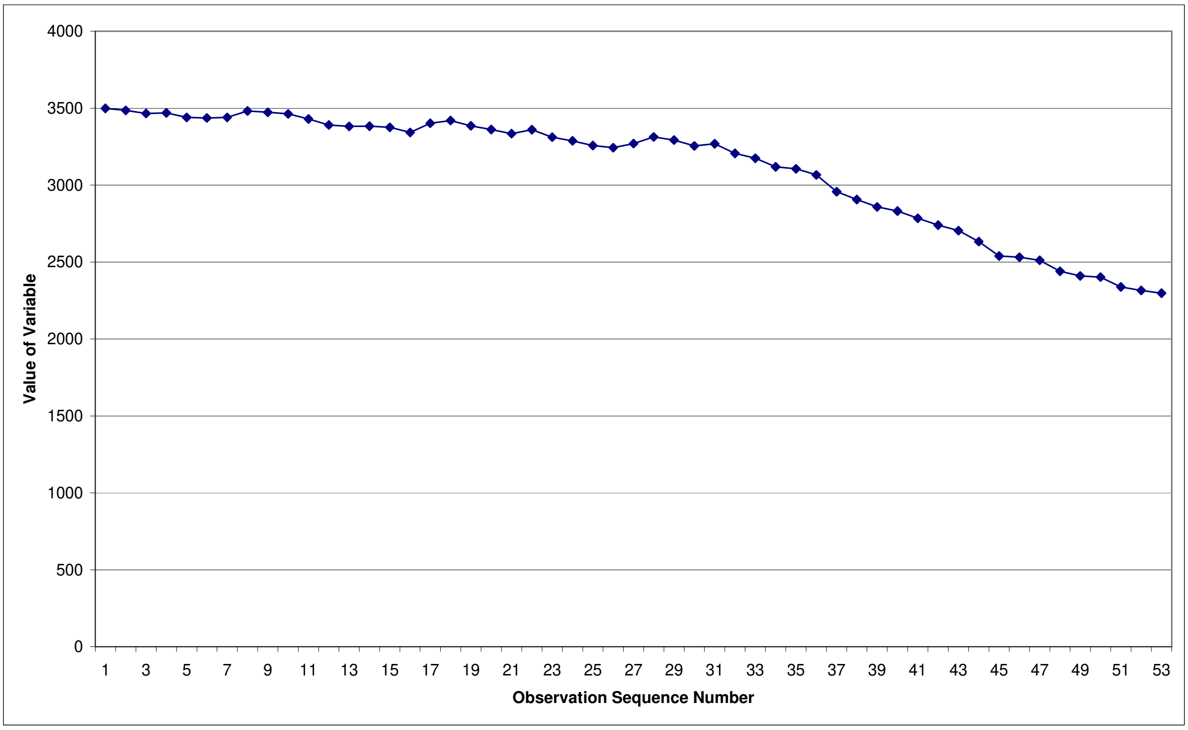
$$MdRAE = median(APE_{i,h,j}),$$

where s is the total number of total number of series forecasted and MdAPE is observation $(s+1)/2$ if s is odd, or the mean of observations $s/2$ when s is even when $APE_{i,h,j}$ observations are ordered by rank. The PB statistic counts the proportion a given method has a forecasting error larger than a relative method:

$$PB_{i,h,j} = \left[\frac{1}{s} \sum_{j=1}^{s} \delta_{i,h,j} \right] * 100$$

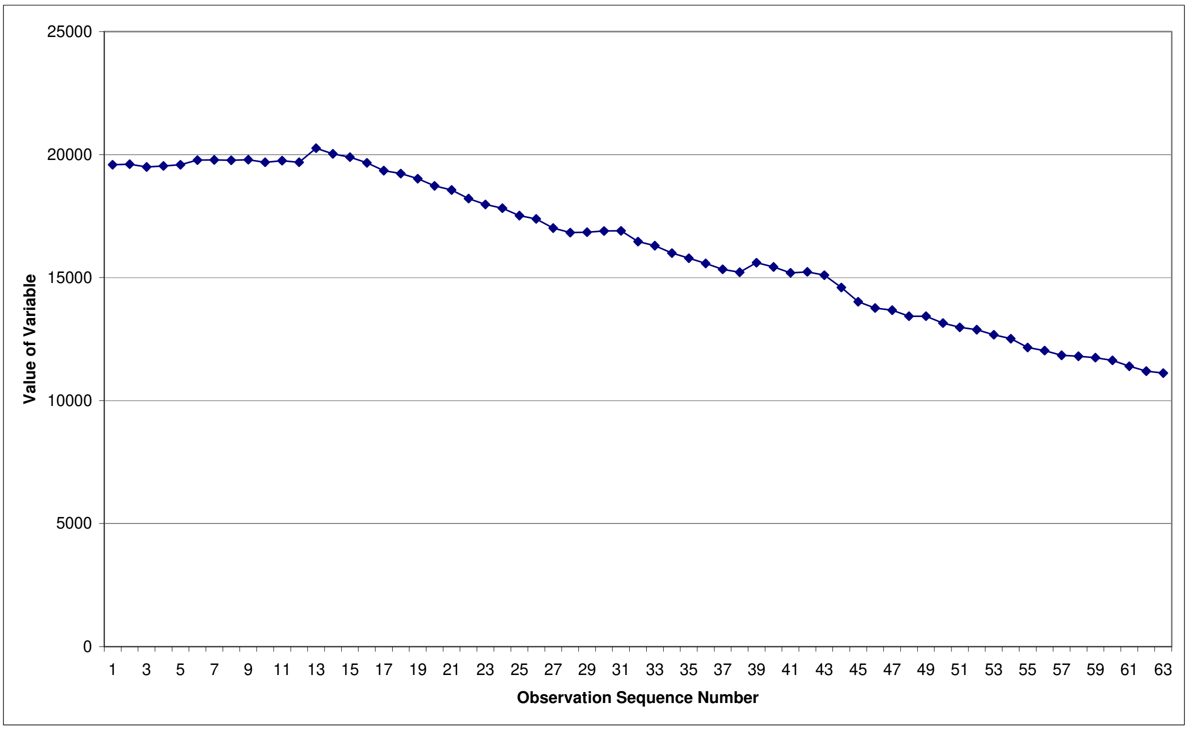
where $\delta_{i,h,j} = \begin{cases} 1 & if \left| F_{i,h,j} - A_{h,j} \right| \prec \left| F_{rw,h,j} - A_{h,j} \right| \\ 0 & otherwise \end{cases}$. $F_{i,h,j}$ is the forecast for method i at horizon h for series j. and $A_{h,j}$ is the actual value at horizon h for series i. $\delta_{i,h,j}$ is a dummy variable that records the proportion of times a particular model i forecasting horizon h for series j has a lower forecast error than the random walk model and s is the total number of series forecasted and $PB_{i,h,j}$ is the proportion (in percentage) method i performs better that the relative method. The relative method applied for comparison in this study is the random walk model. A value of greater than 50% for $PB_{i,h,j}$ indicates the forecasts obtained for a particular forecasting method i is more accurate than the random walk.

Fig. I. Specimen of a M3 Monthly Series



Source. IIF(2005).

Fig. II. Specimen of a M3 'Other' Series



Source. IIF(2005).

Fig. III. Outlier Frequency of M3 Monthly Series

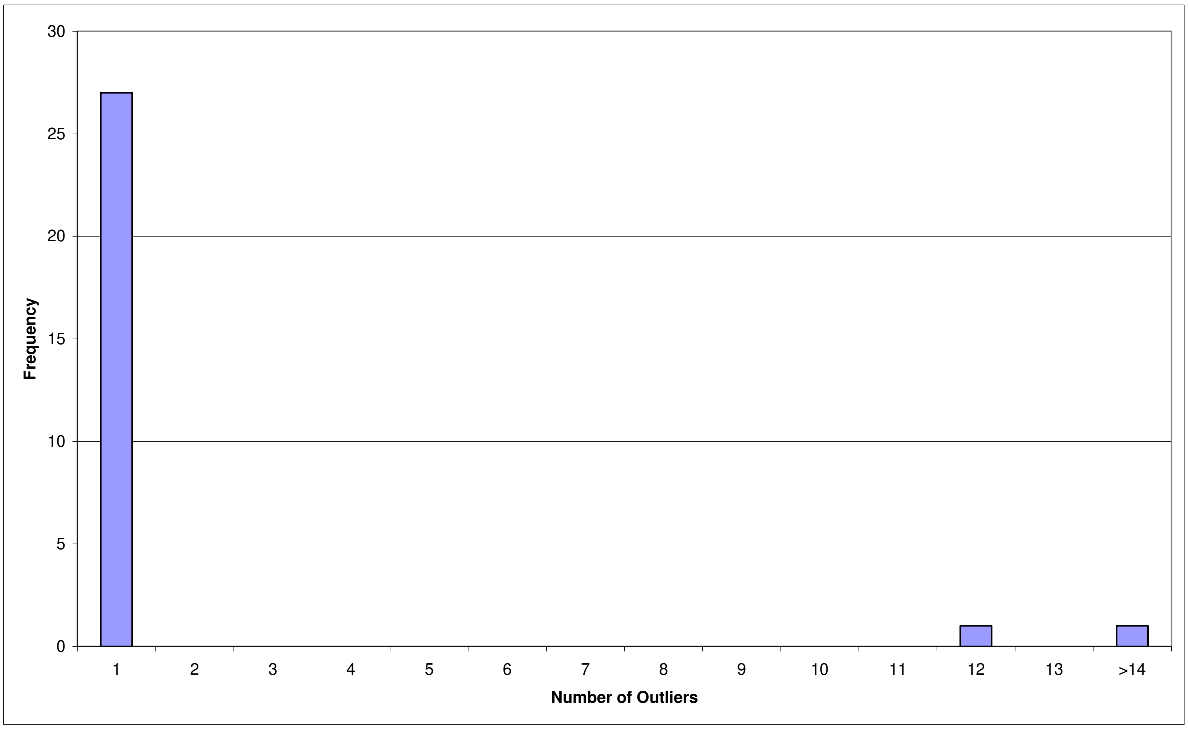


Fig. IV. Outlier Frequency of M3 ‘Other’ Series

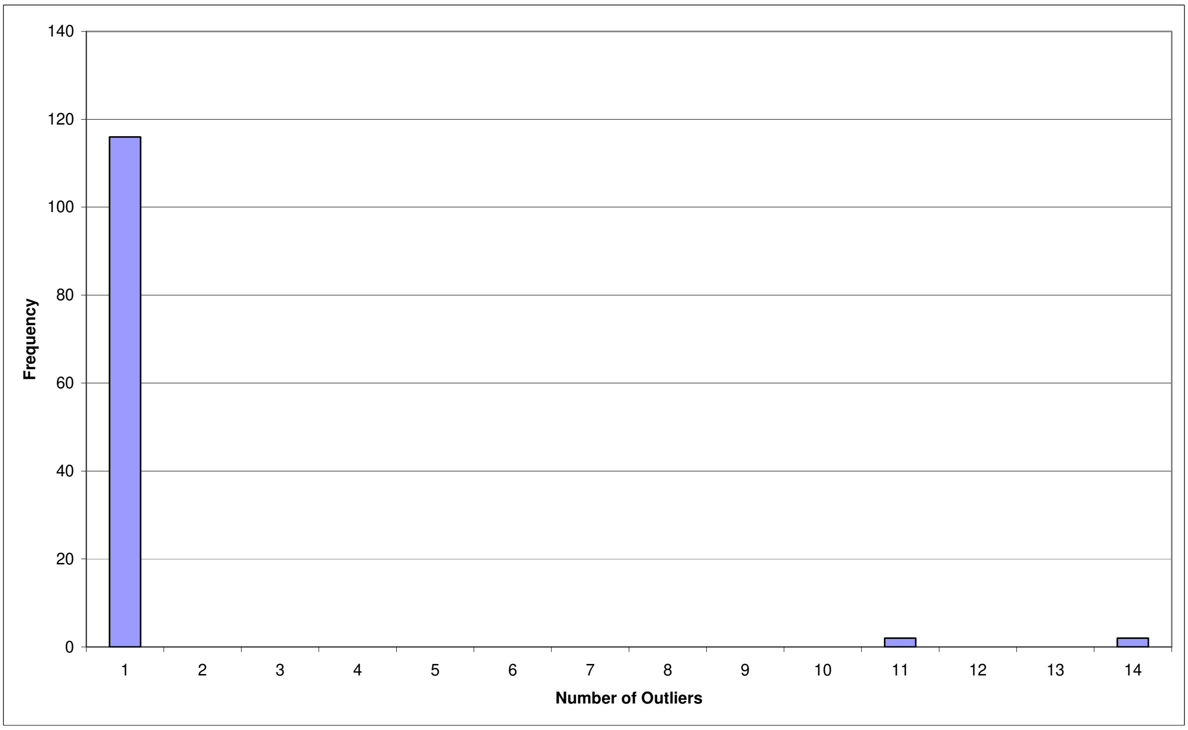


Fig. V. Correlation with Time for M3 Monthly Series

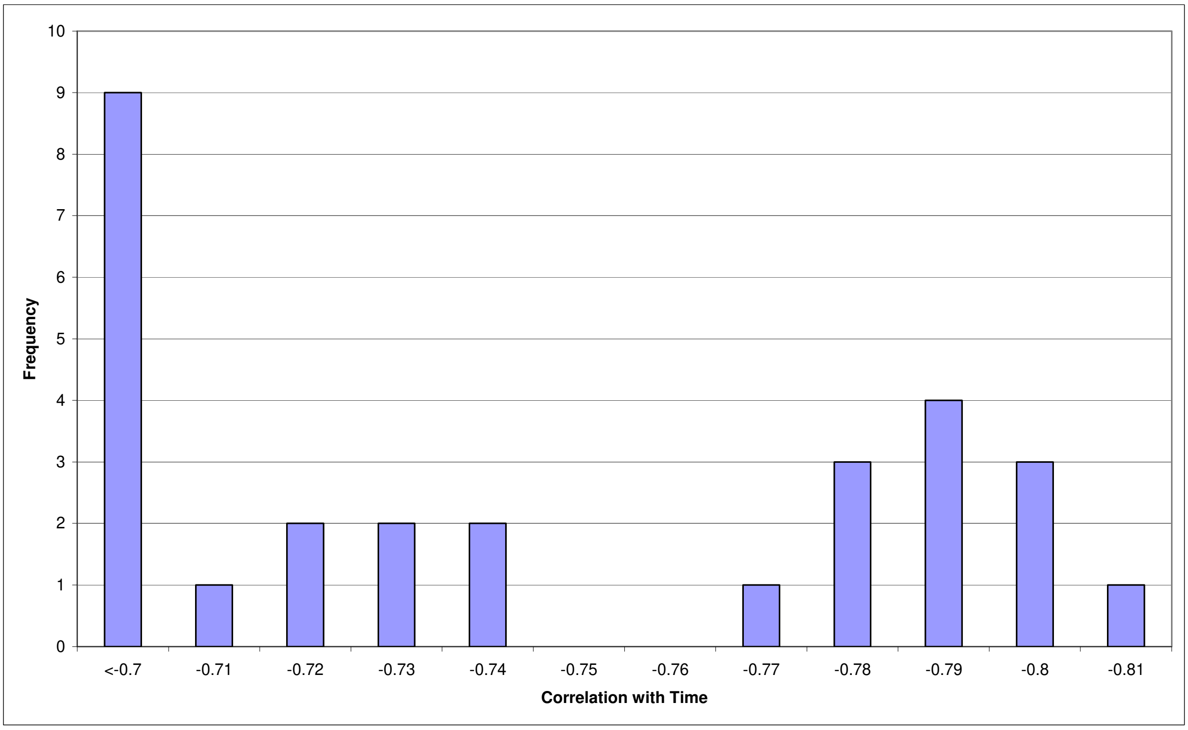


Fig. VI. Correlation with Time for M3 'Other' Series

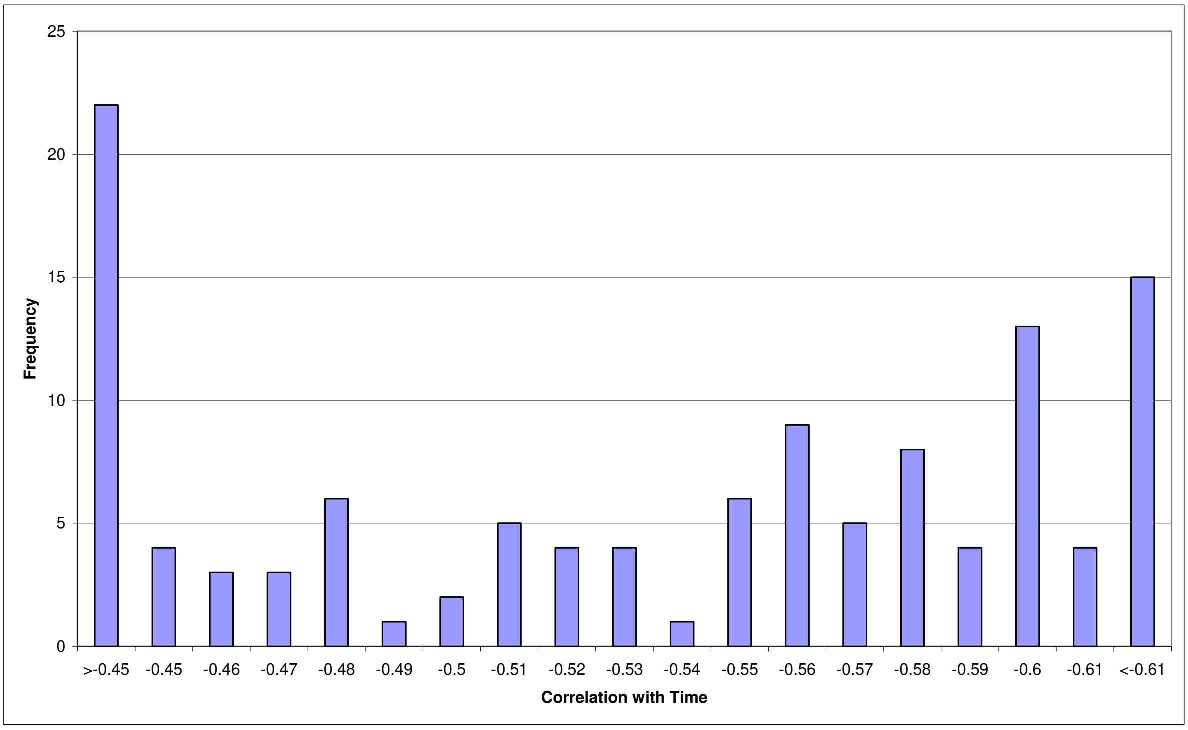


Fig. VII. Variation Explained by the Linear/AR for M3 Monthly Series

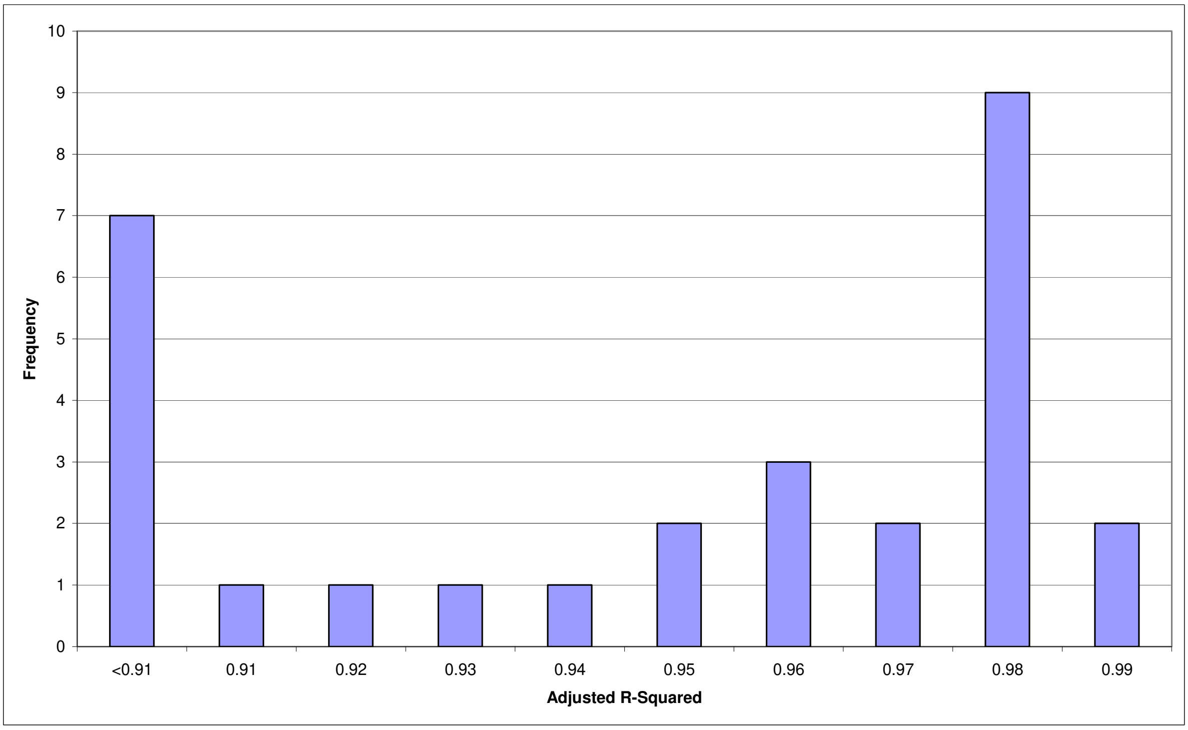


Fig. VIII. Variation Explained by the Linear/AR for M3 ‘Other’ Series

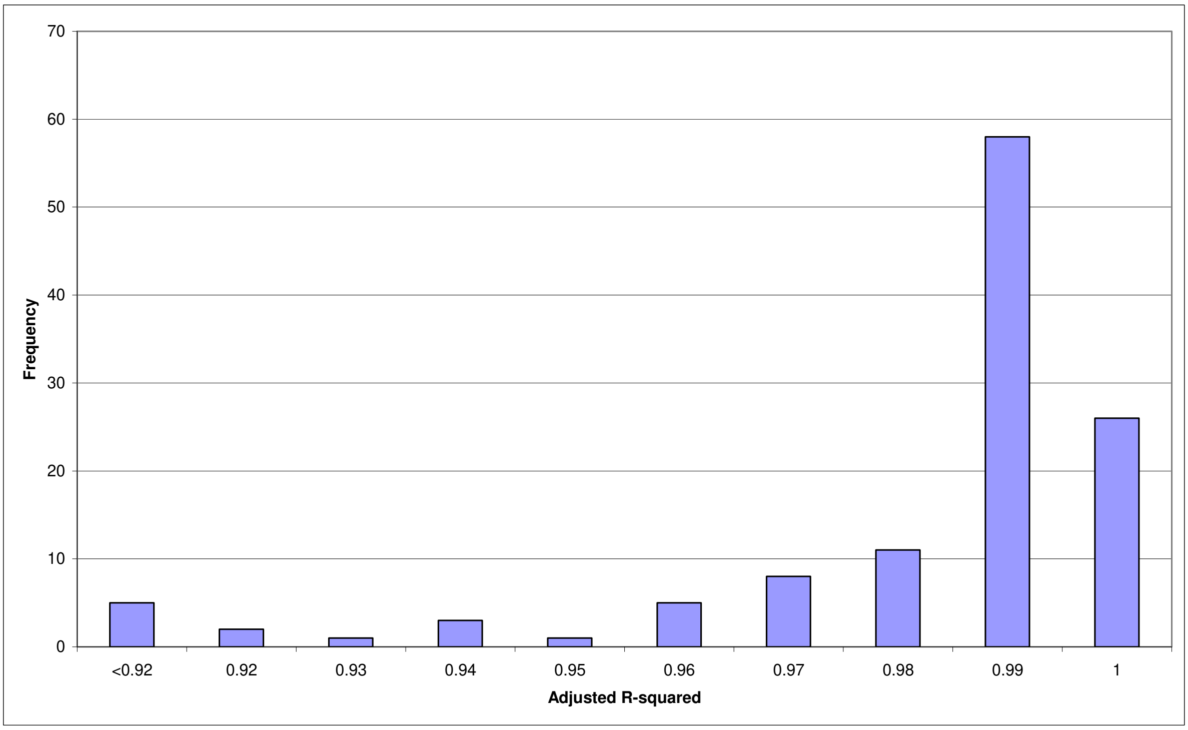


Table I. Summary Statistics for M3 Monthly Series

	Mean	Variance	Skewness	Kurtosis	Minimum	Maximum
Mean	4,953.84	2,337.83	0.88	2.61	2,183.06	10,740.60
Variance	386.80	362.29	2.55	10.37	57.75	1,867.98
Skewness	73.91	36.66	0.07	1.74	11.04	134.92
Kurtosis	-0.82	0.85	1.84	6.35	-1.90	2.13
Turning Points	10.00	3.09	-0.54	2.11	4.00	15.00
Step Changes	0.48	0.87	1.51	3.89	0.00	3.00
Outliers	27.97	6.63	-3.14	11.42	2.00	30.00
CV	0.08	0.08	3.29	14.81	0.02	0.46
Maximum	5,562.57	2,489.95	0.74	2.58	2,384.60	11,855.20
Minimum	4,331.32	2,159.54	0.78	2.26	1,879.20	8,971.50
Runs	16.83	4.98	-0.36	2.21	6.00	24.00
Autocorrelation	0.95	0.07	-3.11	12.36	0.65	0.99

Note: The number of series tested is 29. The summary statistics report represents the summary statistics of each series. Std Dev is the standard deviation; Turing Points and Step Changes are the number of turning points and step changes, respectively, defined in Shah (1997); Outliers is the number of outliers greater than 3 standard deviations; CV is the coefficient of variation; Runs is the number of runs; Autocorrelation is the estimate of autocorrelation of lag 1.

Table II. Summary Statistics of M3 'Other' Series

	Mean	Variance	Skewness	Kurtosis	Minimum	Maximum
Mean	6,748.04	5,988.95	2.75	13.31	1,749.82	42,191.30
Std Dev	921.50	1,315.54	4.37	23.94	109.88	8,988.27
Skewness	102.49	46.78	0.93	5.92	3.70	319.65
Kurtosis	-0.94	0.82	4.84	34.51	-1.90	5.49
Turning Points	12.88	4.62	0.02	2.60	2.00	23.00
Step Changes	0.57	1.00	1.70	4.86	0.00	4.00
Outliers	48.10	6.68	-3.88	17.25	15.00	50.00
CV	0.13	0.06	1.03	4.44	0.03	0.32
Maximum	8,146.49	7,821.10	3.14	15.79	2,127.00	55,794.00
Minimum	5,221.50	4,523.13	2.65	12.67	1,276.00	30,908.00
Runs	17.58	5.46	0.00	2.82	4.00	30.00
Autocorrelation	0.98	0.03	-4.71	27.37	0.81	1.00

Note: The number of series tested is 120. The summary statistics report represents the summary statistics of each series. Std Dev is the standard deviation; Turing Points and Step Changes are the number of turning points and step changes, respectively, defined in Shah (1997); Outliers is the number of outliers greater than 3 standard deviations; CV is the coefficient of variation; Runs is the number of runs; Autocorrelation is the estimate of autocorrelation of lag 1.

Table III. Mean Absolute Percentage Error for M3 Monthly Series

Forecast Method	Forecast Horizon			
	1	6	12	18
RT	0.136	0.182	0.211	0.235
ARIMA	**0.004**	0.077	0.183	0.322
HOLT	0.012	**0.066**	**0.124**	**0.185**
HOLT-D	0.008	0.085	0.397	4.119
HOLT-W	0.016	0.094	0.184	0.292
ARARMA	0.860	0.949	1.104	1.815
SES	0.042	0.119	0.190	0.267
RW	0.013	0.087	0.161	0.236

Note: RT is the robust trend model; ARIMA is the autoregressive integrated moving average model; HOLT is Holt's linear no trend model; Holt-D is Holt's model with exponential smoothing; HOLT-W is the linear no trend Holt-Winters model; ARARMA is a long memory model; SES is the linear no trend simple exponential smoothing model. The RW is the random walk model with no drift. Percentages in bold at a forecast horizon is the models with lowest mean absolute percentage error.

Table IV. Median Absolute Percentage Error for M3 Monthly Series

Forecast Method	Forecast Horizon			
	1	6	12	18
RT	0.116	0.152	0.179	0.199
ARIMA	**0.005**	0.044	0.103	0.137
HOLT	0.006	0.027	0.060	0.083
HOLT-D	0.007	**0.048**	**0.091**	**0.172**
HOLT-W	0.006	0.029	0.075	0.123
ARARMA	0.999	1.006	1.029	1.033
SES	0.033	0.105	0.164	0.232
RW	0.007	0.066	0.120	0.197

Note: RT is the robust trend model; ARIMA is the autoregressive integrated moving average model; HOLT is Holt's linear no trend model; Holt-D is Holt's model with exponential smoothing; HOLT-W is the linear no trend Holt-Winters model; ARARMA is a long memory model; SES is the linear no trend simple exponential smoothing model. The RW is the random walk model with no drift. Percentages in bold at a forecast horizon is the models with lowest median absolute percentage error.

Table V. Percent Better for M3 Monthly Series

Forecast Method	Forecast Horizon			
	1	6	12	18
RT	1.400	10.345	6.897	44.828
ARIMA	**75.862**	72.414	**79.310**	**79.310**
HOLT	55.172	72.414	72.414	72.414
HOLT-D	55.172	**75.862**	**79.310**	75.862
HOLT-W	51.724	72.414	72.414	68.966
ARARMA	13.793	3.448	10.345	10.345
SES	3.448	3.448	6.897	6.897

Note: RT is the robust trend model; ARIMA is the autoregressive integrated moving average model; HOLT is Holt's linear no trend model; Holt-D is Holt's model with exponential smoothing; HOLT-W is the linear no trend Holt-Winters model; ARARMA is a long memory model; SES is the linear no trend simple exponential smoothing model. The RW is not shown as the PB statistic shows the proportion of series for each model at each forecast horizon that is better than the forecasts of the RW model. Percentages in bold at a forecast horizon is the models with highest percentage accuracy.

Table VI. Mean Absolute Percentage Error for M3 'Other' Series

Forecast Method	Forecast Horizon			
	1	4	6	8
RT	0.079	0.084	0.086	0.097
ARIMA	**0.009**	0.034	0.051	0.071
HOLT	0.010	**0.025**	**0.036**	**0.049**
HOLT-D	**0.009**	0.030	0.047	0.070
HOLT-W	0.012	0.032	0.047	0.064
ARARMA	0.978	0.981	0.985	0.994
SES	0.040	0.074	0.096	0.128
RW	0.014	0.047	0.069	0.100

Note: RT is the robust trend model; ARIMA is the autoregressive integrated moving average model; HOLT is Holt's linear no trend model; Holt-D is Holt's model with exponential smoothing; HOLT-W is the linear no trend Holt-Winters model; ARARMA is a long memory model; SES is the linear no trend simple exponential smoothing model. The RW is the random walk model with no drift. Percentages in bold at a forecast horizon is the models with lowest mean absolute percentage error.

Table VII. Median Absolute Percentage Error for M3 'Other' Series

Forecast Method	Forecast Horizon			
	1	4	6	8
RT	0.070	0.073	0.079	0.088
ARIMA	0.006	0.028	0.041	0.058
HOLT	0.008	**0.022**	**0.027**	**0.033**
HOLT-D	**0.006**	**0.022**	0.036	0.050
HOLT-W	0.009	**0.022**	0.033	0.047
ARARMA	1.002	1.002	1.001	1.016
SES	0.037	0.066	0.089	0.119
RW	0.012	0.040	0.062	0.089

Note: RT is the robust trend model; ARIMA is the autoregressive integrated moving average model; HOLT is Holt's linear no trend model; Holt-D is Holt's model with exponential smoothing; HOLT-W is the linear no trend Holt-Winters model; ARARMA is a long memory model; SES is the linear no trend simple exponential smoothing model. The RW is the random walk model with no drift. Percentages in bold at a forecast horizon is the models with lowest median absolute percentage error.

Table VIII. Percent Better for M3 'Other' Series

Forecast Method	Forecast Horizon			
	1	4	6	8
RT	1.667	9.167	17.500	48.333
ARIMA	65.833	67.500	68.333	70.000
HOLT	65.000	75.000	**76.667**	**79.167**
HOLT-D	**69.167**	**75.833**	75.833	76.667
HOLT-W	59.167	70.000	69.167	73.333
ARARMA	1.667	0.833	1.667	1.667
SES	6.667	5.833	5.000	5.833

Note: RT is the robust trend model; ARIMA is the autoregressive integrated moving average model; HOLT is Holt's linear no trend model; Holt-D is Holt's model with exponential smoothing; HOLT-W is the linear no trend Holt-Winters model; ARARMA is a long memory model; SES is the linear no trend simple exponential smoothing model. The RW is not shown as the PB statistic shows the proportion of series for each model at each forecast horizon that is better than the forecasts of the RW model. Percentages in bold at a forecast horizon is the models with highest percentage accuracy.

Fig. IX. Mean Absolute Percentage Error for M3 Monthly Series

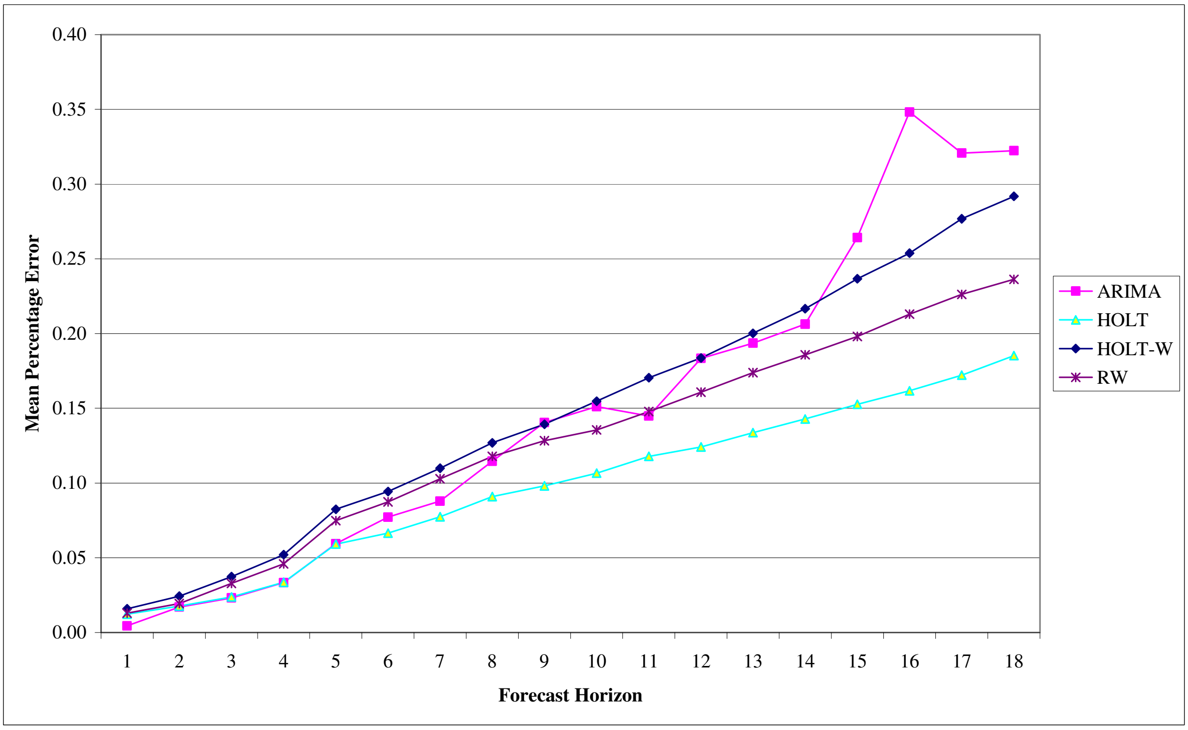


Fig. X. Median Absolute Percentage Error for M3 Monthly Series

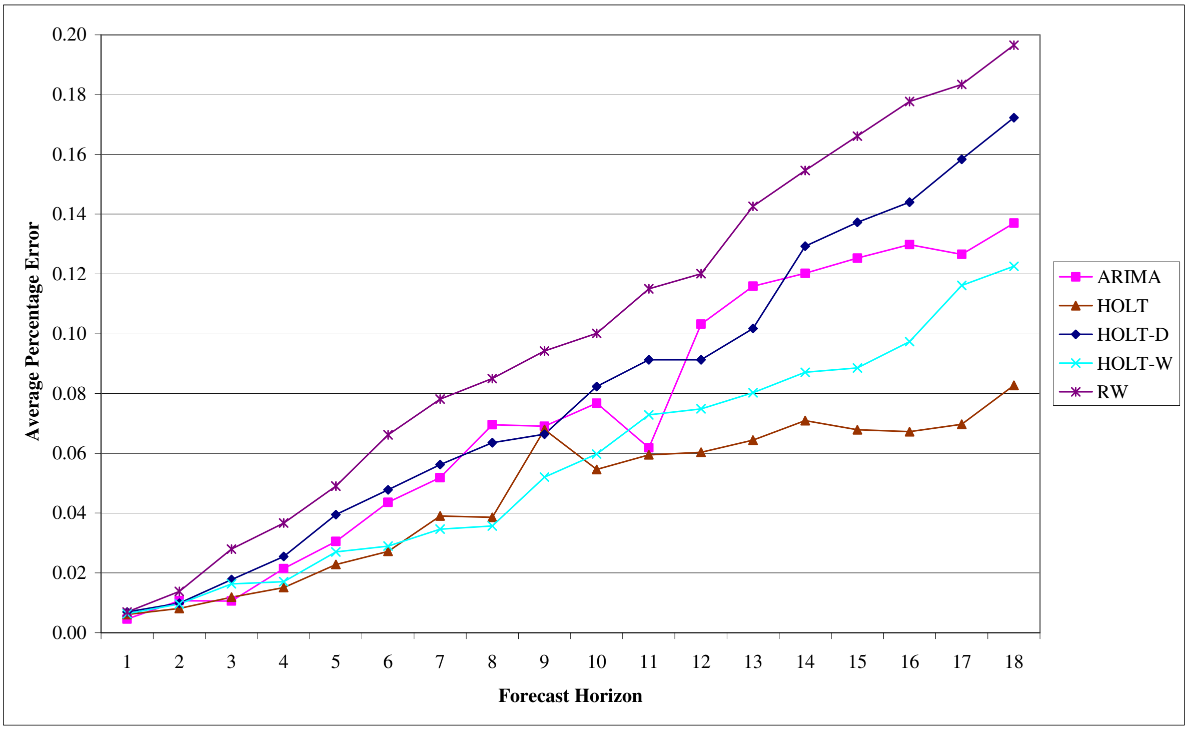


Fig. XI. Percentage Better for M3 Monthly Series

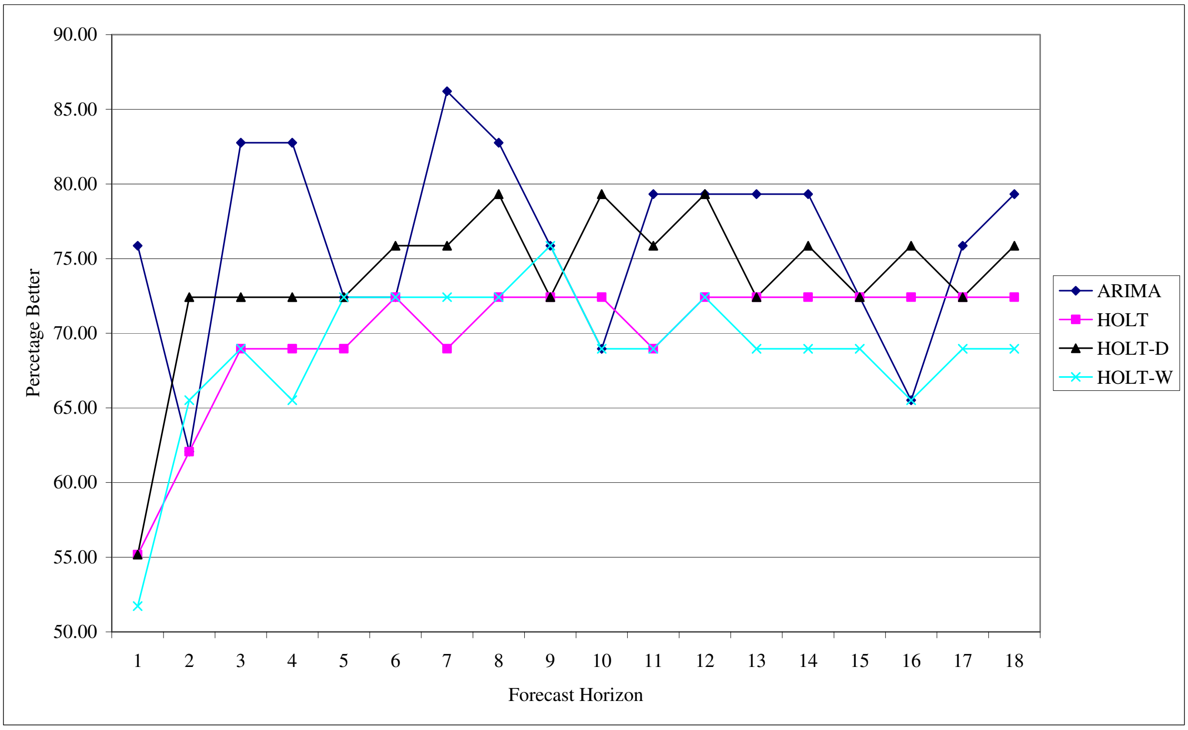


Fig. XII. Mean Absolute Percentage Error for M3 'Other' Series

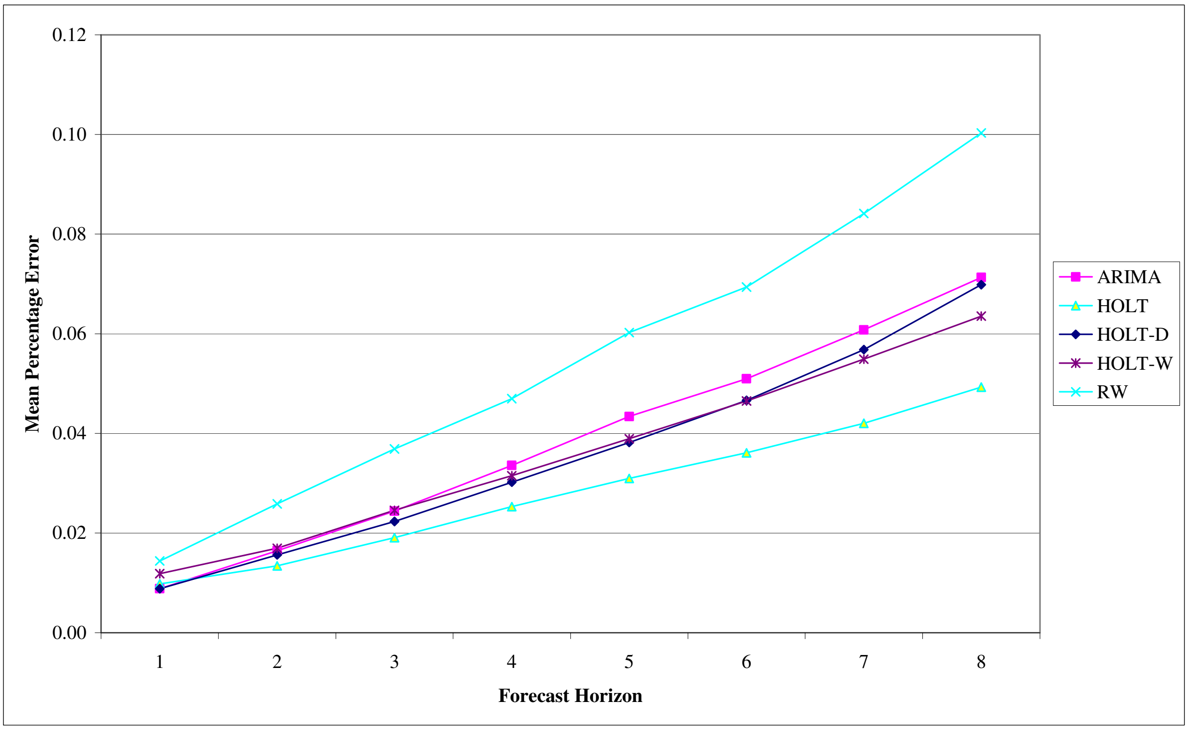


Fig. XIII. Median Absolute Percentage Error for M3 'Other' Series

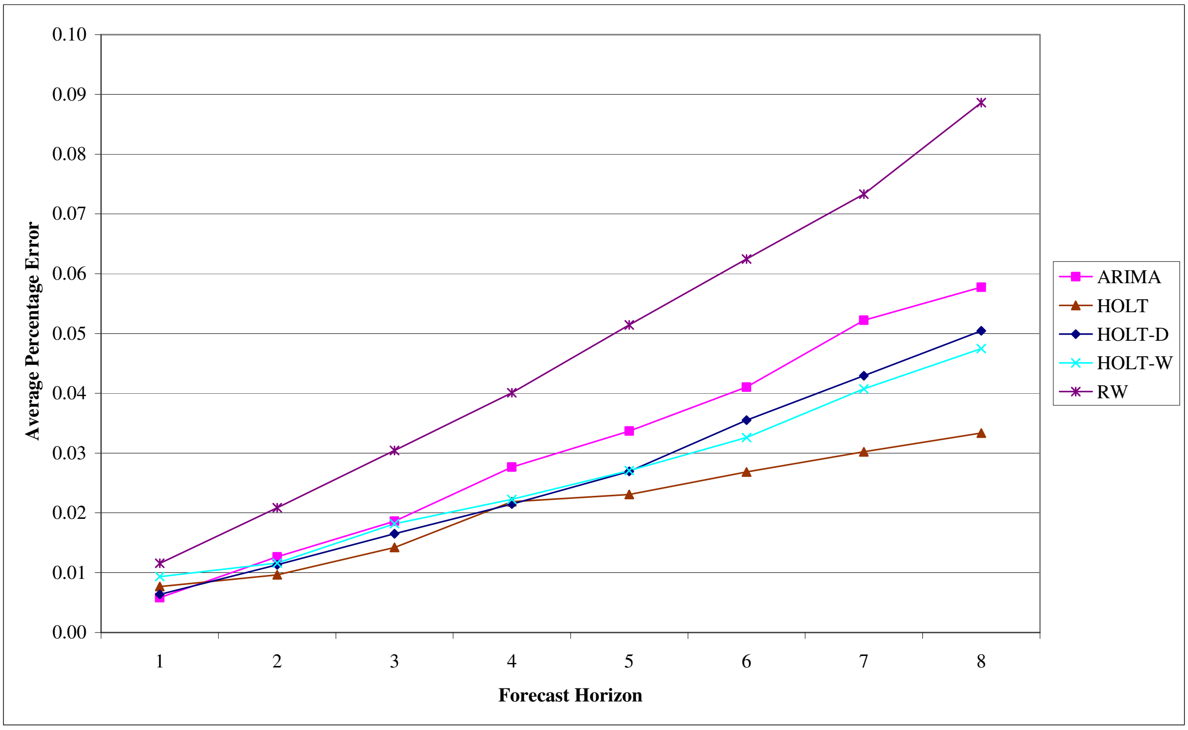


Fig. XIV. Percentage Better for M3 'Other' Series

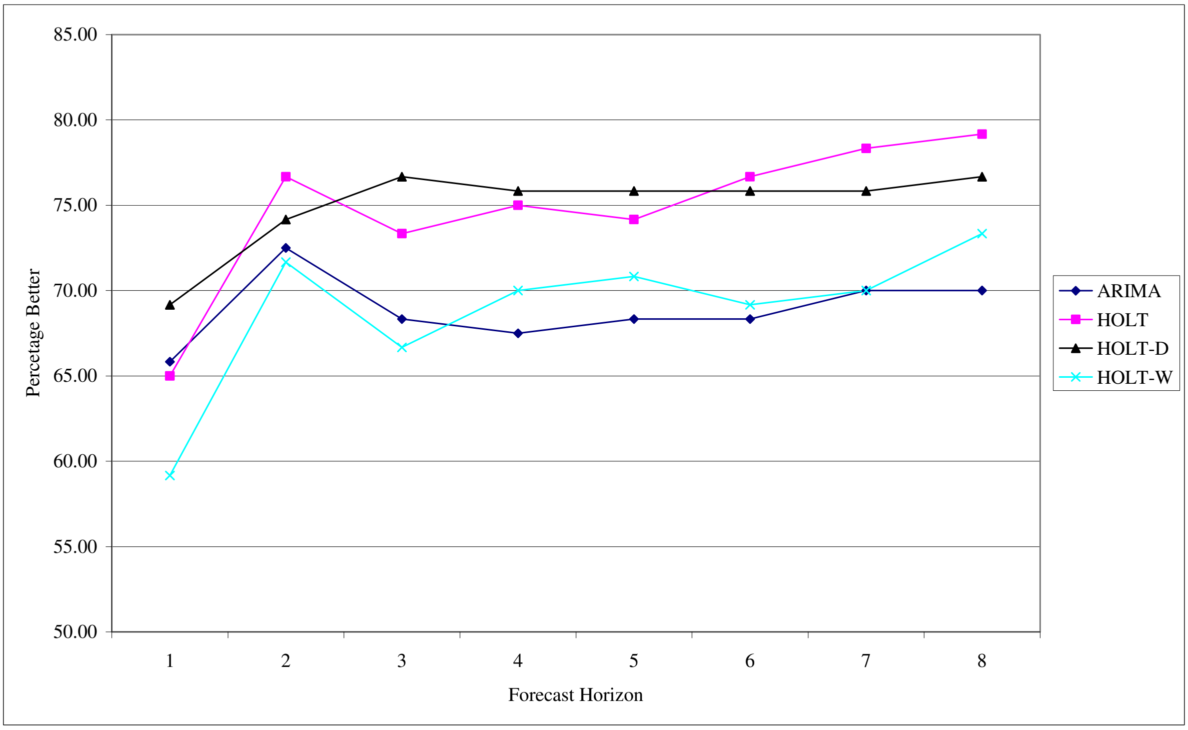


Table IX. Best Models for Forecasting M3 Series by MAPE

M3 Series	Forecast Horizon			
	1	6	12	18
Monthly	ARIMA	HOLT	HOLT	HOLT
Other	ARIMA/ HOLT-D	HOLT	HOLT	HOLT

Note: RT is the robust trend model; ARMA is the autoregressive integrated moving

Table X. Best Models for Forecasting M3 Series by MdAPE

M3 Series	Forecast Horizon			
	1	6	12	18
Monthly	ARIMA	HOLT-D	HOLT-D	HOLT-D
Other	HOLT-D	HOLT/ HOLT-D/ HOLT-W	HOLT	HOLT

Note: RT is the robust trend model; ARMA is the autoregressive integrated moving

Table XI. Best Models for Forecasting M3 Series by PB

M3 Series	Forecast Horizon			
	1	6	12	18
Monthly	ARIMA	HOLT-D	HOLT/ HOLT-D	HOLT-D
Other	HOLT-D	HOLT-D	HOLT	HOLT

Note: RT is the robust trend model; ARMA is the autoregressive integrated moving