CORRELATIONS AMONG VARIOUS FORECAST ACCURACY MEASURES FOR I.I.D. NORMAL TIME SERIES --- A SIMULATION STUDY

Luh Yu (Louie) Ren, University of Houston-Victoria, renl@uhv.edu, 281-275-8832
Peter Ren, University of North Texas, renp@unt.deu, 361-676-5743

ABSTRACT

Linear correlations among various popular forecast accuracy measures are studied through simulation when simple moving average (SMA) methods are applied to identically and independently distributed (i.i.d.) normal time series. This allows us to examine correlations among popular forecasting measures.

Keywords: MAD, RMSE, MAPE, U_2, MASE

INTRODUCTION

When a time series has no trend, no seasonality, and no cycle, that is, under the most ideal scenario of following an identical and independent distribution (i.i.d.), the simple moving average method for forecasting will be applied.

Different measures have been proposed to evaluate forecasts in the past few decades, including Armstrong & Collope, 1992, Chatfield, 1998, Collopy, F. and Amstrong, J. 2000, Fildes. R. & Makridakis, S., 1995, Goodwin, P. and Lawton, R., 1999, Hyndman, R. and Koehler, A. B. , 2006, Koehler, A. B., & Murphree, E. S., 1988, Makridakis, S., & Hibon, M., 1979, Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E., & Winkler, R., 1982, Makridakis, S., Chatfield, C., Hibon, M., Lawrence, M., Mills, T., Ord, K., Simmons, L. F., 1993, Newbold, P., 1983, Reid, D. J., 1975, Ren, L. and Glasure, Y., 2009, and Theil, H., 1966. The most commonly used absolute measures include Mean Absolute Deviation (MAD), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Geometric Mean Square Error (GMSE). Popular relative measures include Mean Absolute Percentage Error (MAPE), Median Absolute Percentage Error (MdAPE), Symmetric (or Unbiased) Mean Absolute Error (sMAPE), and Symmetric (or Unbiased) Median Absolute Error (sMdAPE). Finally, relative measures, such as U_2 proposed by Their in 1966, Relative Mean Absolute Error (RMAE), Relative Median Absolute Error (RMdAE), and scaled relative measure proposed by Hyndman and Koehler in 2006 have also been developed.

Very few statistical relationships of forecast accuracy measures have been established due to the dependency of the error terms and ratios of the measures. A well-known example of this relationship can be found in Geary (1936). His paper shows that the ratio $\frac{1}{\sqrt{n}}\frac{\sum|e_i-\bar{e}|}{\sqrt{\sum(e_i-\bar{e})^2}}$ is asymptotically equal to $\sqrt{\frac{2}{\pi}}$, or simply MAD=RMSE=0.80σ, when e_i's are i.i.d. from a normal probability distribution and the sample size n approaches infinity.

In practical applications, the sample size n seldom approaches infinity. In addition, instead of using the overall average $\bar{e}$ as the forecaster, the simple moving averages (SMA) are used. The probability distributions as well as the correlations among various forecast accuracy measures are also difficult to find. For instance, it is very difficult to find, mathematically or statistically, the correlations between the Mean Absolute Deviation (MAD) and the Root Mean Squared Errors (RMSE), when the SMA method with moving period of 3 is used across various forecasting horizons. Therefore, a simulation study is conducted in this paper. This study focuses on five (5) forecast measures in particular: MAD, RMSE, MAPE, U_2, and Scaled MSE.

As described in Ren and Glasure (2009), MAPE is sensitive to i.i.d. time series with different coefficient of variation (c.v.). Therefore, in order to examine the correlations among various forecast measures, a simulation study for errors from SMA is conducted under the most ideal situation of i.i.d. normal probability distribution with c.v.'s at 0.5.

DATA ANALYSIS

In this study, five (5) forecast accuracy measures, MAD, RMSE, MAPE, U_2, and Scaled MSE (MASE) are studied are assessed. Definitions for these forecast accuracy measures are listed in Appendix A.

20,000 random data are simulated from normal distributions with a mean of one and a standard deviation of 0.5 (i.e., with moderate coefficients of variation (c.v.) of 0.5). The data is then grouped into 1,000 groups with 20 observations each.

The first nine observations in each group are treated as historical observations, and the tenth to twentieth observations are treated as the future 11 observations. Moving Average Methods with moving periods of 1, 3, 5, 7, 9 are applied to historical observations and their forecasts compared with the first future observation (the tenth observation). An error term, A_{10}-F_{10} is calculated for the first future observation.

The eleventh observation is now included into the new historical group. Moving Average Methods with moving periods of 1, 3, 5, 7, and 9, are applied to the most current nine historical observations counted back from the tenth observation (i.e., the first observation in the old historical data group is eliminated), and their forecasts are compared with the "new" first future observation (the eleventh observation). An error term, A_{11}-F_{11} is calculated. This process is repeated to calculate errors up to an including the eleventh error term, A_{20}-F_{20}.

The above process is then applied to a total of 1,000 simulated time series set (each with 20 observations). The MAD, RMSE, MAPE, U_2, and Scaled MAE (MASE) are obtained.

After acquiring the 1,000 independent observations from each forecast accuracy measure and each SMA method, the correlation coefficients are found and listed in Table 1 for forecasting horizons k= 3, 5, 7, 9, and 11.

Table 1: Correlation Coefficients for Various Forecast Accuracy Measures at Different Forecast Horizons

		MAD					RMSE				
		k=3	k=5	k=7	k=9	k=11	k=3	k=5	k=7	k=9	k=11
RMSE	MA(1)	0.981	0.975	0.970	0.967	0.967	-----	-----	-----	-----	-----
	MA(3)	0.969	0.959	0.958	0.954	0.954	-----	-----	-----	-----	-----
	MA(5)	0.968	0.959	0.956	0.952	0.952	-----	-----	-----	-----	-----
	MA(7)	0.968	0.957	0.953	0.952	0.952	-----	-----	-----	-----	-----
	MA(9)	0.969	0.955	0.951	0.949	0.949	-----	-----	-----	-----	-----
MAPE	MA(1)	0.170	0.196	0.187	0.184	0.196	0.159	0.171	0.170	0.175	0.185
	MA(3)	0.156	0.172	0.177	0.183	0.176	0.154	0.167	0.173	0.175	0.169
	MA(5)	0.150	0.154	0.152	0.161	0.154	0.150	0.165	0.168	0.175	0.172
	MA(7)	0.154	0.159	0.162	0.160	0.151	0.157	0.169	0.173	0.173	0.169
	MA(9)	0.170	0.173	0.177	0.178	0.167	0.169	0.177	0.183	0.186	0.178
U_2	MA(1)	0.499	0.577	0.581	0.601	0.638	0.483	0.547	0.553	0.56	0.598
	MA(3)	0.354	0.478	0.476	0.520	0.541	0.327	0.437	0.437	0.471	0.492
	MA(5)	0.402	0.449	0.484	0.485	0.514	0.377	0.402	0.431	0.423	0.447
	MA(7)	0.294	0.434	0.473	0.527	0.544	0.269	0.391	0.406	0.465	0.485
	MA(9)	0.289	0.382	0.463	0.510	0.519	0.248	0.319	0.399	0.436	0.440
MASE	MA(1)	0.775	0.709	0.665	0.637	0.595	0.764	0.693	0.647	0.615	0.571
	MA(3)	0.753	0.680	0.631	0.593	0.551	0.722	0.647	0.603	0.558	0.515
	MA(5)	0.749	0.663	0.626	0.579	0.538	0.723	0.643	0.602	0.553	0.510
	MA(7)	0.747	0.665	0.620	0.575	0.536	0.726	0.645	0.598	0.551	0.511
	MA(9)	0.750	0.666	0.614	0.572	0.529	0.730	0.641	0.589	0.546	0.499
		MAPE					U_2				
		k=3	k=5	k=7	k=9	k=11	k=3	k=5	k=7	k=9	k=11
U_2	MA(1)	0.145	0.195	0.198	0.165	0.191	-----	-----	-----	-----	-----
	MA(3)	0.093	0.130	0.105	0.113	0.135	-----	-----	-----	-----	-----
	MA(5)	0.083	0.112	0.077	0.068	0.100	-----	-----	-----	-----	-----
	MA(7)	0.063	0.124	0.121	0.112	0.127	-----	-----	-----	-----	-----
	MA(9)	0.051	0.127	0.158	0.132	0.135	-----	-----	-----	-----	-----
MASE	MA(1)	0.167	0.182	0.171	0.129	0.135	0.405	0.406	0.376	0.363	0.366
	MA(3)	0.157	0.165	0.157	0.125	0.117	0.217	0.324	0.299	0.301	0.294
	MA(5)	0.151	0.145	0.141	0.11	0.105	0.312	0.286	0.292	0.263	0.255
	MA(7)	0.157	0.154	0.147	0.112	0.108	0.237	0.297	0.281	0.272	0.250
	MA(9)	0.174	0.164	0.159	0.123	0.116	0.218	0.245	0.294	0.272	0.240

CONCLUSION

Table 1 shows that MAD and RMSE are highly linearly correlated, as well as being moderately linearly correlated with MASE. Therefore, any of these measures can be used interchangeably as a forecast accuracy measure under the stated conditions. However, MAPE and U_2 are not recommended. This result is consistent with findings from previous research. Also, non-linear relationships among those forecast accuracy measures or correlations from non-normal probability distributions will be further studied in separate research.

REFERENCES

Armstrong, J. S., & Collopy, F. (1992). Error measures for generalizing about forecasting methods: empirical comparisons (with discussion). International Journal of Forecasting, 8, 69-80, 99-111.

Chatfield, C. (1998). Apples, oranges, and mean squared error. *International Journal of Forecasting* 4: 515-518.

Collopy, F., and Amstrong, J. (2000). Another error measure for selection of the best forecasting method: the unbiased absolute percentage error. http://hops.wharton.upenn.edu/forecast/paperpdf/armstrong-unbiasedAPE.pdf. 1-5

Fildes. R., & Makridakis, S. (1995). The impact of empirical accuracy studies on time series analysis and forecasting. International Statistical Review, 63, 289-308.

Geary, R. (1936). Moments of the ratio of the mean deviation to the standard deviation for normal samples. *Biometrika* 28: 295-307.

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

Hyndman, R., and Koehler, A. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting* 22: 679-688.

Koehler, A. B., & Murphree, E. S. (1988). A comparison of results from state space forecasting with forecasts from the Makridakis competition. International Journal of Forecasting, 4, 45-55.

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

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

Makridakis, S., Chatfield, C., Hibon, M., Lawrence, M., Mills, T., Ord, K., & Simmons, L. F. (1993). The M-2 Competition: a real-time judgmentally based forecasting study. International Journal of Forecasting, 9, 5-23.

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

Newbold, P. (1983). The competition to end all competitions. Journal of Forecasting, 2, 276-279.

Reid, D. J. (1975). A review of short term projection techniques. In: Gordon, H. D. (Ed.), Practical aspects of forecasting, Operational Research Society, London, pp. 8-25.

Ren, L., & Glasure, Y. (2009). Applicability of the Revised Mean Absolute Percentage Errors (MAPE) Approach to Some Popular Normal and Non-normal Independent Time Series. *International Advances In Economic Research*, 15:4, 409-420.

Theil, H. (1966). *Applied economic forecasting*. Chicago: Rand McNally.