ELSEVIER

International Journal of Forecasting 15 (1999) 405–408

international journal of forecasting

www.elsevier.com/locate/ijforecast

On the asymmetry of the symmetric MAPE

Paul Goodwin*, Richard Lawton

Faculty of Computer Studies and Mathematics, University of the West of England, Coldharbour Lane, Frenchay, Bristol, BS16 1QY, UK

Abstract

Several authors have suggested that the use of the mean absolute percentage error (MAPE) as a measure of forecast accuracy should be avoided because they argue it treats forecast errors above the actual observation differently from those below this value. To counter this, the use of a symmetric (or modified) MAPE has been proposed. This paper shows that, in its treatment of negative and positive errors, the proposed modification is far from symmetric, particularly where these errors have large absolute values. It also shows that, under some circumstances, a non-monotonic relationship can occur between the symmetric MAPE and the absolute forecast errors.

Keywords: Forecasting; Error measures

1. Introduction

When the performance of forecasting methods needs to be compared across different time series, accuracy measures like the mean squared error (MSE) and the mean absolute error (MAE) are inappropriate. This is because there can often be major variations in the scale of the observations between series so that a few series with large values can dominate the comparisons (Chatfield, 1988). In these circumstances, unit free measures need to be employed and the mean absolute percentage error (MAPE) is probably the most widely used measure of this kind. The MAPE is defined below for forecasts made for periods 1 to N of a single series.

The forecast error at time $t = e_t = A_t - F_t$ where A_t = the actual observation at time t and F_t = the forecast made for period t.

$$\text{The percentage error} = \frac{A_t - F_t}{A_t} \times 100$$

so that the absolute percentage error for period t

$$(\text{APE}_t) = \left| \frac{A_t - F_t}{A_t} \right| \times 100$$

$$\text{and the MAPE} = \sum_{t=1}^{N} \text{APE}_t / N$$

Despite its widespread use, the MAPE has several disadvantages (Armstrong & Collopy, 1992; Makridakis, 1993). In particular, Makridakis has argued that the MAPE is asymmetric in that 'equal errors above the actual value result in a greater APE than those below the actual value'. Similarly, Armstrong and Collopy argued that 'the MAPE . . . puts a heavier penalty on forecasts that exceed the actual than those that are less than the actual. For example,

*Corresponding author. Tel.: +44-117-965-6261; fax: +44-117-976-3860.

E-mail address: paul.goodwin@uwe.ac.uk (P. Goodwin)

0169-2070/99/$ – see front matter
PII: S0169-2070(99)00007-2

the MAPE is bounded on the low side by an error of 100%, but there is no bound on the high side'. To correct this, Makridakis then proposed the use of a modified APE which involves dividing the absolute error by the *average* of the actual observation and the forecast. The formula for this measure is shown below.

$$\text{Modified APE for period } t = \left| \frac{A_t - F_t}{(A_t + F_t)/2} \right| \times 100$$

This measure has since been referred to as the SAPE (smoothed absolute percentage error) (O'Connor et al., 1997) and, when averaged over a set of forecasts, the SMAPE (symmetric mean absolute percentage error) (O'Connor and Lawrence, 1998). To demonstrate the measure, Makridakis compared the APE of (i) an actual of 150 and a forecast of 100 and (ii) an actual of 100 and a forecast of 150. In the first case, he showed that the APE is 33.33% and in the second case, 50%, while the modified APE gives a value of 40% in both cases. Makridakis suggested, albeit in the basis of preliminary empirical evidence, that the modification of the MAPE 'may be the most appropriate way to meet theoretical and practical concerns, and do so in a simple and meaningful way'. Recently, an anonymous referee suggested to one of us that the measure should have been used in a study in preference to the median absolute percentage error (MdAPE). (The MdAPE had been chosen for use in this study because it is not distorted by extreme APEs that may arise from the occasional small actual value and because it performed well under many conditions in an empirical study by Armstrong and Collopy (1992).)

2. Problems with the modified APE

As we show below, although the modified APE is symmetric when A_t and F_t are *interchanged*, it in fact creates a new problem of asymmetry which is more likely to be of practical concern than the problem resulting from the interchange. Indeed, the conventional APE does *not* treat single errors above the actual value any differently from those below it. If the actual value is 100 units, errors of -10 and $+10$ units both result in an APE of 10%. The modified APE *does* treat them differently. For example, the errors of -10 and $+10$ units, given above, would result in modified APEs of 18.18% and 22.2%, respectively. Fig. 1 shows the modified APEs for percentage errors ranging from -100% to $+100$% *when the actual value* (A_t) *is 100 units*. It

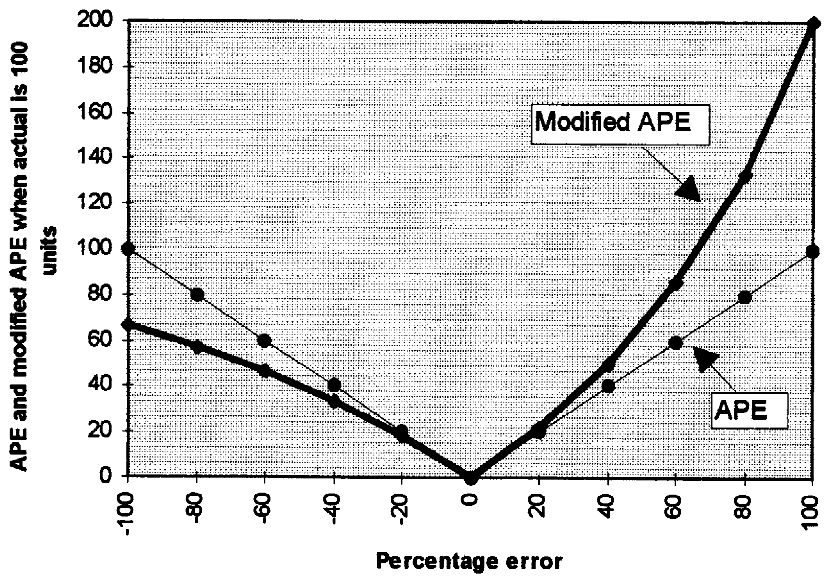

Fig. 1. Modified APE is asymmetric relative to the APE.

can clearly be seen that positive and negative errors are treated in a quite different way. Not only will an error of $-x\%$ result in a different modified APE from an error of $+x\%$, but the rate of increase in the modified APE as the absolute percentage error increases also depends on the sign of the original error. While small positive and negative errors may have similar modified APEs, for large errors the differences can be considerable. For example, the diagram shows that when the forecast error is $+100\%$ the modified APE is three times higher than when the error is -100%. Furthermore, as long as both actual and forecast have positive values, the modified APE for both positive and negative errors has an upper bound of 200%. However, while this bound is reached when $F_t = 0$ for positive errors, for negative errors it is only approached as the error at time t, e_t, tends to minus infinity. This can be seen by rearranging the formula for the modified APE to obtain:

$$\text{Modified APE} = 100|e_t(A_t - e_t/2)| = \frac{200}{\dfrac{2A_t}{|e_t|} + 1} \quad \text{when } e_t < 0$$

so that the modified APE approaches 200% as e_t tends to minus infinity.

When the *forecasts and actuals are of opposite sign* (e.g., a forecast of a positive temperature and an actual observation that turns out to be negative) very large modified APEs can occur. This happens when the *absolute* values of the forecast and actual are close so that the denominator of the modified APE is very small. Indeed, when the absolute values of forecast and actual are the same, but the signs of the original values are different, the measure is undefined (though, of course, there is an identical problem with the conventional APE when $A_t = 0$). Furthermore, if either:

1. the actual value is positive, but the forecast is negative and $e_t > 2A_t$ or
2. the actual value is negative, but the forecast is positive and $e_t < -2A_t$

then increases in the absolute errors actually lead to a reduction in the modified APE. For example, given an actual value of -200 units, a forecast of 300 units gives a modified APE of 1000%, while the less accurate forecast of 400 units gives a modified APE of 600%.

However, our main concern about the modified APE is the lack of symmetry in the way it treats positive and negative errors. This can have serious implications when these errors are large and we can see no justification for it in general practical applications. Of course, the pattern shown in Fig. 1 might fortuitously approximate the forecast error loss function in a specific practical problem, but even then issues would need to be addressed about the boundary between forecasting and decision making (Goodwin, 1996). It might be argued that the use of the modified APE is justified because it lessens the effect of occasional small actual values. For example, a forecast of 11 units and an actual value of 1 unit gives a conventional APE of 1000%, which in the conventional MAPE would tend to distort the resulting mean. In contrast, the modified APE here would only be 166.7%. However, in these circumstances, the use of the median of the conventional APEs (i.e. the MdAPE), or any sensible trimming procedure, will avoid the distortion anyway.

3. Conclusion

The selection of appropriate error measures in forecasting is always a problem because, as Mathews and Diamantopoulos (1994) point out, no single measure gives an unambiguous indication of forecasting performance, while the use of multiple measures makes comparisons between forecasting methods difficult and unwieldy. Nevertheless, we would caution against the use of the modified (or symmetric) MAPE on the grounds that it treats large positive and negative errors very differently.

References

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

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

Goodwin, P. (1996). Statistical correction of judgmental point forecast and decisions. *Omega: International Journal of Management Science 24*, 551–559.

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

Mathews, B. P., & Diamantopoulos, A. (1994). Towards a taxonomy of forecast error measures: a factor-comparative investigation of forecast error dimensions. *Journal of Forecasting 13*, 409–416.

O'Connor, M., & Lawrence, M. (1998). Judgemental forecasting and the use of available information. In: Wright, G., & Goodwin, P. (Eds.), Forecasting With Judgment, Wiley, Chichester, pp. 65–90.

O'Connor, M., Remus, W., & Griggs, K. (1997). Going up-going down: how good are people at forecasting trends and changes in trends? *Journal of Forecasting 16*, 165–176.

Biographies: Paul GOODWIN is Principal Lecturer in Operational Research at the University of the West of England. His research interests focus on the role of judgment in forecasting and decision making. He has published articles in a number of academic journals including the *International Journal of Forecasting*, the *Journal of Forecasting*, *Omega* and the *Journal of Behavioral Decision Making*.

Richard LAWTON is Head of the School of Operational Research at the University of the West of England. His research interests focus on the performance and behaviour of exponential smoothing methods.