International Journal of Forecasting 32 (2016) 669–679

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

A new metric of absolute percentage error for intermittent demand forecasts

Sungil Kim [a], Heeyoung Kim [b,*]

[a] Data Analytics Lab, R&D Center, Samsung SDS, Seoul, Republic of Korea

[b] Department of Industrial and Systems Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea

ARTICLE INFO

Keywords:
Accuracy measure
Forecast evaluation
Intermittent demand
MAPE

ABSTRACT

The mean absolute percentage error (MAPE) is one of the most widely used measures of forecast accuracy, due to its advantages of scale-independency and interpretability. However, MAPE has the significant disadvantage that it produces infinite or undefined values for zero or close-to-zero actual values. In order to address this issue in MAPE, we propose a new measure of forecast accuracy called the *mean arctangent absolute percentage error* (*MAAPE*). MAAPE has been developed through looking at MAPE from a different angle. In essence, MAAPE is a *slope as an angle*, while MAPE is a *slope as a ratio*, considering a triangle with adjacent and opposite sides that are equal to an actual value and the difference between the actual and forecast values, respectively. MAAPE inherently preserves the philosophy of MAPE, overcoming the problem of division by zero by using bounded influences for outliers in a fundamental manner through considering the ratio as an angle instead of a slope. The theoretical properties of MAAPE are investigated, and the practical advantages are demonstrated using both simulated and real-life data.

1. Introduction

The mean absolute percentage error (MAPE) is one of the most popular measures of the forecast accuracy. It is recommended in most textbooks (e.g., Bowerman, O'Connell, & Koehler, 2004; Hanke & Reitsch, 1995), and was used as the primary measure in the M-competition (Makridakis et al., 1982). MAPE is the average of absolute percentage errors (APE). Let A_t and F_t denote the actual and forecast values at data point t, respectively. Then, MAPE is defined as:

$$\text{MAPE} = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{A_t - F_t}{A_t}\right|, \tag{1.1}$$

where N is the number of data points. To be more rigorous, Eq. (1.1) should be multiplied by 100, but this is omitted in this paper for ease of presentation without loss of generality. MAPE is scale-independent and easy to interpret, which makes it popular with industry practitioners (Byrne, 2012).

However, MAPE has a significant disadvantage: it produces infinite or undefined values when the actual values are zero or close to zero, which is a common occurrence in some fields. If the actual values are very small (usually less than one), MAPE yields extremely large percentage errors (outliers), while zero actual values result in infinite MAPEs. In practice, data with numerous zero values are observed in various areas, such as retailing, biology, and finance, among others. For the area of retailing, Fig. 1 (Makridakis, Wheelwright, & Hyndman, 1998) illustrates typical intermittent sales data. Many zero sales occur dur-

* Corresponding author. Tel.: +82 42 350 3125; fax: +82 42 350 3110.
E-mail address: heeyoungkim@kaist.ac.kr (H. Kim).

http://dx.doi.org/10.1016/j.ijforecast.2015.12.003

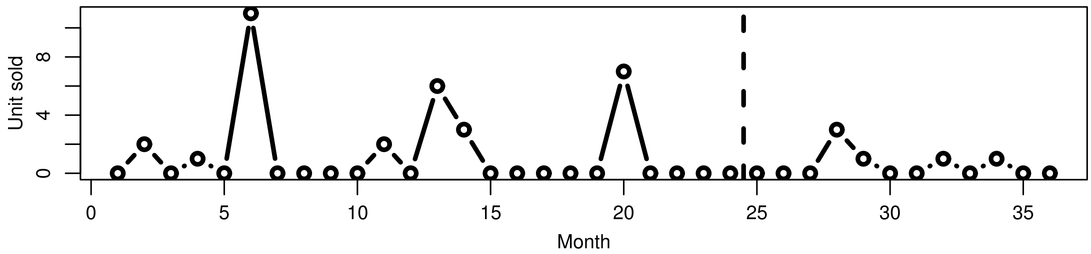

Fig. 1. Three years of monthly sales of a lubricant product sold in large containers. Data source: 'Product C' from Makridakis et al. (1998, Ch. 1). The vertical dashed line indicates the end of the data used for fitting and the start of the data used for out-of-sample forecasting.

ing the time periods considered, and this leads to infinite or undefined MAPEs.

There have been attempts to resolve this problem by excluding outliers that have actual values less of than one or APE values greater than the MAPE plus three standard deviations (Makridakis, 1993). However, this approach is only an arbitrary adjustment, and leads to another question, namely how the outliers can be removed. Moreover, the exclusion of outliers might distort the information provided, particularly when the data involve numerous small actual values. Several alternative measures have been proposed to address this issue. The symmetric mean absolute percentage error (sMAPE), proposed by Makridakis (1993), is a modified MAPE in which the divisor is half of the sum of the actual and forecast values. Another measure, the mean absolute scaled error (MASE), was proposed by Hyndman and Koehler (2006). The MASE is obtained by scaling the forecast error based on the in-sample mean absolute error using the naïve (random walk) forecast method, and can overcome the problem of the MAPE generating infinite or undefined values. Similarly, Kolassa and Schütz (2007) proposed that the mean absolute error be scaled by the in-sample mean of the series (MAE/Mean ratio) in order to overcome the problem of division by zero.

While these alternative measures resolve the MAPE's issue with outliers, the original MAPE remains the preferred method of business forecasters and practitioners, due to both its popularity in the forecasting literature and its intuitive interpretation as an *absolute percentage error*. Therefore, this paper proposes an alternative measure that has the same interpretation as an *absolute percentage error*, but can overcome the MAPE's disadvantage of generating infinite values for zero actual values.

Even though this paper focuses on MAPE, it is worth reviewing the other accuracy measures used in the literature as well. In general, accuracy measures can be split into two groups: scale-dependent measures and scale-independent measures. As the group names indicate, the scale-dependent measures are measures for which the scale depends on the scale of the data. The mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE), and median absolute error (MdAE) all belong to this category. These measures are useful when comparing different forecasting methods that are applied to data with the same scale, but should not be used when comparing forecasts for series that are on different scales (Chatfield, 1988; Fildes & Makridakis, 1988). In that situation, scale-independent measures are more appropriate. Being scale-independent has been considered to be a key characteristic for a good measure (Makridakis, 1993). The aforementioned MAPE, sMAPE, MASE, and the MAE/Mean ratio are examples of scale-independent measures.

There have been various attempts in the literature to make scale-dependent measures scale-independent by dividing the forecast error by the error obtained from a benchmark forecasting method (e.g., a random walk). The resulting measure is referred to as a relative error. The mean relative absolute error (MRAE), median relative absolute error (MdRAE), and geometric mean relative absolute error (GMRAE) all belong to this category. Even though Armstrong and Collopy (1992) recommended the use of relative absolute errors, particularly the GMRAE and MdRAE, these measures have the issue of potentially involving division by zero. In order to overcome this difficulty, Armstrong and Collopy (1992) recommended that extreme values be trimmed; however, this increases both the complexity and the arbitrariness of the calculation, as the amount of trimming must be specified.

Relative measures are another type of scale-independent measure. Relative measures are similar to relative errors, except that relative measures are based on the values of measures instead of errors. For example, the relative MSE (RelMSE) is given by the MSE divided by MSE_b, where MSE_b denotes the MSE from a benchmark method. Similar relative measures can be defined using RMSE, MAE, MdAE, MAPE, and so on. A log-transformed RelMSE, i.e., log(RelMSE), has also been proposed, in order to impose symmetrical penalties on the errors (Thompson, 1990). When the benchmark method is a random walk and the forecasts are all one-step forecasts, the relative RMSE is Theil's U statistic (Theil, 1966, Ch. 2), which is one of the most popular relative measures. However, Theil's U statistic has the disadvantages that its interpretation is difficult and outliers can easily distort the comparisons because it does not have an upper bound (Makridakis & Hibon, 1979). In general, relative measures can be highly problematic when the divisor is zero. For a more in-depth review of other accuracy measures, refer to Hyndman and Koehler (2006), who provide an extensive discussion of various measures of forecast accuracy, and Hyndman (2006), particularly for measures for intermittent demand.

The remainder of this paper is organized as follows. In Section 2, MAPE is investigated from a different angle, with a new measure called MAAPE being proposed as a result. The behavior and theoretical properties of the proposed measure are then investigated in Section 3. In Section 4, we further explore the bias aspect of MAAPE in comparison