The Accuracy Measurement of Stock Price Numerical Prediction

Samsul Amar¹, Andi Sudiarso², Muhammad Kusumawan Herliansyah³

¹Department of Industrial Engineering, Faculty of Engineering, Universitas Trunojoyo Madura. (Doctoral Student of Department of Mechanical and industrial Engineering, Faculty of Engineering, Universitas Gadjah Mada) Yogyakarta, Indonesia

², ³Department of Mechanical and Industrial Engineering, Faculty of Engineering, Universitas Gadjah Mada, Yogyakarta, Indonesia

Email: sams_amar@yahoo.com¹, a.sudiarso@ugm.ac.id², herliansyah@ugm.ac.id³

Abstract. Stock market prediction is both attracting and challenging. The successful of stock price prediction will give financial benefit, therefore, it has attracted many researchers and practitioners since long time ago. One of stock prediction performance measurement is the accuracy. There are many methods for measuring the accuracy, ranging from simple to sophisticated mathematical formulation. A method may be suitable for some type of data of condition to be forecasted while other methods may be suitable for other conditions. This article describes some of those methods. MAPE is the most popular accuracy measure for forecasting since it is intuitive, easy to be interpreted and can be applied to measure the forecasting accuracy for both individual item and across item groups. However, the MAPE has shortcomings. First, the value of MAPE will be undefined if one or more of the actual data are zero. Second, MAPE is very sensitive to outlier data. Third, equal errors above the actual value result in a greater percentage error than those below the actual value. Finally, using the MAPE for forecast methods comparison will lead to a systematically under-forecast result. Even with its shortcomings, MAPE is preferred for the forecasting accuracy method due to its applicability, interpretability and reliability. Thus, it is better to look for ways of correcting the drawbacks of MAPE rather than searching for alternative measures which are less desirable than MAPE. This paper proposes a modified MAPE using moving average of the actual data (MAPEMA) to reduce the drawbacks of MAPE. The result shows that the use of MAPEMA can reduce the problems of using MAPE.

1. Introduction

Stock market prediction is both attracting and challenging. By its nature, stock market is complex, nonlinear, and volatile, and therefore predicting stock market is a difficult task. The successful of stock market prediction will give financial benefit, therefore, it has attracted many researchers and practitioners since long time ago. Many methods have been developed and applied for stock market prediction. The methods can be divided into technical analysis, fundamental analysis, and technological methods [1]. In technical analysis, the assumption is that the pattern of stock price or movement is follow the pattern of the past, and therefore, predictable. The fundamental analysis assumes that the stock price, sooner or later, will go to its intrinsic value. Therefore, to get an abnormal return, investor should find the stocks those are having the price below their intrinsic value. Technological methods apply statistical or artificial intelligence to make the prediction.

The type of stock market prediction can be divide into two main categories: numerical prediction and direction prediction. An example of the numerical prediction is stock price prediction where the result of prediction, is the price, in currency unit, of some stocks. Other examples of numerical prediction are the prediction of stock return, volume and index. Stock return is proportion or percentage of the stock price increase (or decrease) to its previous price. The second type of stock market prediction tries to predict the movement direction of the price or index in the future, whether it will up or down. Sometimes, the direction is categorized into several levels, for example, it may be divided into big down, down, stable, up, and high up.

One of stock prediction performance measurement is the accuracy. In prediction, accuracy, or goodness of fit, is the degree to which the result of the prediction conforms the actual value. Accuracy of is an important thing in the stock price prediction Obviously, high accuracy is more preferred than low accuracy. Low accuracy can lead to high uncertainty in the return or even worse can result in losses. Researchers and practitioners are in the race to find prediction methods those produce high accuracy constantly and applicable. There are many methods for measuring the accuracy, ranging from simple to sophisticated mathematical formulation. A method may be suitable for some type of data of condition to be forecasted while other methods may be suitable for others condition. A good accuracy measurement method should be able to measure the accuracy...
consistently without any significant bias, and also easy to be implemented and interpreted. This article describes some of those methods and the application of those methods for the stock price prediction for numerical prediction. The result of numerical prediction is numeric value. As stated before, examples of the numerical prediction are stock price, stock return, volume and index prediction. The error of the prediction is simply the different between the actual value and the forecast. Accuracy can be described as how close the forecast with the actual value. The accuracy measurement for numerical prediction can be divided into three type:

- scale-dependent measurement,
- percentage base measurement, and
- relative error base measurement.

Some measurement tools from each type will be described in this section. For the rest of the paper, these notations will be used. Suppose there are n period of data and a forecast have been done using a particular method. Error for each period t can be calculated based on the actual data and the forecast result for each period.

\[ e_t = Y_t - F_t \] (1)

Where:
- \( e_t \): forecasting error of the period t,
- \( Y_t \): the actual observation of period t,
- \( F_t \): the forecast of period t

2. Methods

a. Scaled-dependent measurement

The value of scale-dependent measurement is dependent by the scale of the data, however, it is difficult to compare between one to another forecasting of different data or series using this method. For example, the forecasting of a company stock cannot be compared to another company since the price are different. However, this method may be used for comparing two different forecasting method with the same data or series. This method may also be used for the data that has typical scale such as stock return since the stock return has the same scale for all stocks. Below are some methods in this type of measurement.

**Average of Errors (E)**

\[ E = \frac{\sum_{t=1}^{n} e_t}{n} \] (2)

**Mean Absolute Error (MAE) or Mean Absolute Deviation (MAD)**

\[ MAE = MAD = \frac{\sum_{t=1}^{n} |e_t|}{n} \] (3)

**Median Absolute Error (MdAE)**

\[ MdAE = Median(|e_t|) \] (4)

**Sum Squared Error (SSE)**

\[ SSE = \sum_{t=1}^{n} e_t^2 \] (5)

**Mean Squared Error (MSE) or Mean Squared Prediction Error (MSPE)**

\[ MSE = MSPE = \frac{\sum_{t=1}^{n} e_t^2}{n} \] (6)

**Root Mean Squared Error (RMSE)**

\[ \sqrt{RMSE} = \sqrt{\frac{\sum_{t=1}^{n} e_t^2}{n}} \] (7)

In the average error (E), the direction of the error, can eliminate each other, therefore, average error cannot be used for measuring the accuracy performance since a big error will have the same E as a small error if the direction balance in the same. However, this method can be used to measure the error direction tendency. A forecast with positive E tend to be over-estimate while negative value of E tends to be under-estimate.

MAE or MAD has a simple formula making this measure easy to be implemented. However, it is highly affected by outliers and still scale-dependent. This measure can be applied for comparing prediction methods for return prediction, since the scale of return, in term of proportion or percentage, is tend to same for all stock. Moreover, this method can be interpreted easily for return prediction. MdAE is less affected by outliers but still scale dependent and tend to avoid the unbalance density in the calculation.

SSE, MSE and RMSE are the variance of square error. These are popular methods, mostly because of their theoretical relevance in statistical modelling. The bigger errors are disliked and therefore given more weight, the
error itself. However, they are more sensitive to outliers than MAE or MdAE, which has led some authors to recommend against their use in forecast accuracy evaluation. Rankings of methods based on the RMSE were highly unreliable [2]–[4].

b. Percentage base measurement

These measures are not dependent of the data scale, and therefore, suitable for comparing forecast result of different data. However, these measures have the disadvantage of being infinite or undefined if $Y_t = 0$ for any $t$ in the period of interest, and probably having large value when $Y_t$ is close to zero [3]. Below are some methods in this type of measurement.

**Mean Absolute Percentage Error (MAPE)**

$$MAPE = \frac{\sum_{t=1}^{n} |\frac{e_t}{Y_t}|}{n} \times 100\%$$  \hspace{1cm} (8)

**Median Absolute Percentage Error (MdAPE)**

$$MdAPE = Median\left(\frac{|e_t|}{|Y_t|} \times 100\%\right)$$  \hspace{1cm} (9)

**Root Mean Square Percentage Error (RMSPE)**

$$RMSPE = \sqrt{\frac{\sum_{t=1}^{n} (\frac{e_t}{Y_t})^2}{n}}$$  \hspace{1cm} (10)

**Root Median Square Percentage Error (RMdSPE)**

$$RMdSPE = \sqrt{Median\left(\frac{(\frac{e_t}{Y_t})^2}{|Y_t|}\right)}$$  \hspace{1cm} (11)

**Symmetric Mean Absolute Percentage Error (sMAPE)**

$$sMAPE = \frac{\sum_{t=1}^{n} |\frac{2\cdot e_t}{Y_t + F_t}|}{n} \times 100\%$$  \hspace{1cm} (12)

**Symmetric Median Absolute Percentage Error (sMdAPE)**

$$sMdAPE = Median\left(\frac{2\cdot |e_t|}{Y_t + F_t} \times 100\%\right)$$  \hspace{1cm} (13)

**Median Absolute Prediction Error as a percentage of the Standard deviation (MdAPES)**

$$MdAPES = Median\left(\frac{|e_t|}{SD} \times 100\%\right)$$  \hspace{1cm} (14)

Where: $SD$ = the standard deviation of $Y_t$

MAPE is the most popular accuracy measure for forecasting [5]. The reasons are it is intuitive, easy to be interpreted and can be applied to measure or compare accuracy for both individual item and across item groups [6]. However, the MAPE has shortcomings. Some researchers has revealed the disadvantages of using MAPE [3], [6]–[9]. First, the value of MAPE, and also all of its variance, will be undefined if one or more of the actual demands are zero, and it explodes if there is actual data which are very small compared to the forecast, and if the value of the actual data is close to zero. Second, outlier data can make a significant change of MAPE. Third, equal errors above the actual value result in a greater APE (Absolute Percentage Error) than those below the actual value [9]. For example, when the actual value is 150 and the forecast is 100 the APE is 33.33%, however, when the actual is 100 and the forecast 150 the APE is 50%. Forth, using the MAPE for forecast methods comparison will lead to a systematically under-forecast result, especially for series that fluctuate widely [6]. The later problem is illustrated nicely by [6] using the rolling dice problem. The dice roll could represent natural random fluctuation for actual data without trend, seasonality, or causal factors to influence to the data. Suppose a dice is rolled and forecaster should predict the result for each rolling. The expected value in this case will be 3.5 and that is indeed the best forecast in this situation. However, using MAPE, forecast of 2 will lead to better performance than forecast with 3.5. This problem arise as the low denominator will produce bigger APE than high denominator for the same error, thus, the undervalue forecast will tend to have the smaller APE than the higher forecast.
Some researchers suggest the use of median instead of mean for calculating percentage error while others suggest the use of relative error as describe in the next sub-section. [2] suggest the use of MdAPE to select among the forecasting methods when many series of data are available while [10] proposes the MdAPES. The use of MdAPE can reduce the effect of outlier data, however, it cannot reduce the other drawbacks. Moreover, medians are not relative measures and they are not recommended for general use [9]. The use of square percentage error does not eliminate any disadvantage of the MAPE, instead it increases the severe from the outlier data. Even with its shortcomings, Makridakis [9] argues that MAPE is the only choice for the forecasting accuracy method due to its applicability, interpretability and reliability. Thus, he suggests that it is better to look for ways of correcting the drawbacks of MAPE rather than searching for alternative measures which are less desirable than MAPE. He suggests the use of sMAPE. However, [11] show that sMAPE is also produce an asymmetric result. This paper proposes a modified MAPE using moving average of the actual data (MAPEMA) to reduce the drawbacks of MAPE.

3. Result

3.1. Relative Base Measurement

Another way of scaling is to divide each forecasting error by the error obtained using another standard method of forecasting. Let denote \( rt = \frac{et}{et^*} \) as the relative error, where \( et^* \) is the forecast error obtained from the benchmark method. Usually, the benchmark method is the random walk, or naïve, where \( Ft \) is equal to actual data the last period before the forecasting period, \( Y_{t-1} \) [2].

**Mean Relative Absolute Error (MRAE)**

\[
MRAE = \frac{1}{n} \sum_{t=1}^{n} |rt|
\]

**Median Relative Absolute Error (MdRAE)**

\[
MdRAE = \text{Median}(|rt|)
\]

**Geometric Mean Relative Absolute Error (GMRAE)**

\[
GMRAE = \sqrt[n]{\prod_{t=1}^{n} |rt|}
\]

**Percent Better**

\[
\text{Percent Better} = \frac{\text{count}_i \{ |et| \leq |et^*| \}}{n} \times 100%\]

**Relative Mean Absolute Error (RelMAE) or Cumulative Relative Absolute Error (CumRAE)**

\[
\text{RelMAE} = \text{CumRAE} = \frac{MAE}{MAE^*}
\]

Where: \( MAE^* = MAE \) of the benchmark

**Theil’s U2 (U2)**

\[
U2 = \frac{RMSE}{RMSE^*}
\]

Where: \( RMSE^* = RMSE \) of the benchmark

**Mean Absolute Scaled Error (MASE)**

\[
MASE = \frac{1}{n} \sum_{t=1}^{n} |\frac{et}{MAE^*}|
\]

The use of relative error can reduce the effect of low value denominator of MAPE. [2] suggest the use of GMRAE to calibrate the parameters of a given model, and recommend MdRAE for suggest selecting among forecasting methods when using a small number of time series. However, they warned that median error measures, and also percent better measure, are not sensitive because further improvements in forecasting that series produce no change when summarizing across series. Moreover, percent better does not consider the magnitude of error. Big error and small error have the same value if they have the same case where they are better than the benchmark forecasting method. MASE is recommended by [3] and it is proved work well for the dice rolling problem [6]. MASE is also supported by [12] since it has the property of Diebold and Mariano properties [13]. The use of MAE* as the denominator in RelMAE and MASE make the error is equally weighted. In term of stock price prediction, it is unfavorable. For example, the forecasting error of \$ 0.2 will be much different between when the price of a stock is \$ 1 and when the
price of the stock is at $10. The same condition may occur in demand forecasting. Therefore, this paper suggests the use of actual data moving average as the denominator as will be explained later.

3.2. MEAN ABSOLUTE PERCENTAGE ERROR AS PERCENTAGE OF MOVING AVERAGE DATA (MAPEMA)

As stated before, MAPE is the most popular accuracy measure method due to its applicability, interpretability and reliability. However, there are some drawbacks of MAPE that make this method widely criticized. The main cause of the problem in MAPE is the denominator. MAPE use the actual data $Y_t$ as the denominator. When $Y_t$ is zero, or close to zero, it can ruin the accuracy. Moreover, the use of $Y_t$ as the denominator will lead to the under-forecast tendency since the low actual data will result the bigger percentage error and cause the forecast go closer to the low value than it should be. The proposed method, MAPEMA, uses moving average of the actual data as the denominator, see (22).

$$MAPEMA = \frac{\sum_{t=1}^{n} \left| \frac{e_t}{\sum_{i=t-M}^{t-1} Y_i} \right|}{n} \times 100\%$$ (22)

Where: $M = moving\ average\ period$

Table 1. shows simulated data to illustrate the problem of using MAPE and how MAPEMA can overcome the problem. Forecasting result 1 (F1) is obtained using linear regression or least square error method while forecasting result 2 (F2) is obtained from F1 subtracted by 0.2. It is obvious that F1 is better forecast than F2. Fig.1 clearly shows that F1 is visually represent the trend of actual data better than F2. However, if MAPE is used for the accuracy measure, F2 has the lower MAPE than F1 and indicate that F2 is a better forecast compared with F1. This problem arise because the low value of actual data as denominator will produce bigger absolute percentage error than high denominator for the same error. As a result, the undervalue forecast will tend to have the smaller MAPE than the higher forecast. This problem can be reduced if the denominator is the moving average of the actual data as in MAPEMA. Table I shows that F1 get the better (lower) MAPEMA than F2 and indicate that F1 is a better forecast than F2 as it should be.

| t | Act | F1 | F2 | APE F1 | APE F2 | APEMA F1 | APEMA F2 |
|---|-----|----|----|--------|--------|----------|----------|
| 1 | 0.91 |     |    |        |        |          |          |
| 2 | 0.94 |     |    |        |        |          |          |
| 3 | 1.36 | 1.60| 1.40| 17.46  | 2.78   | 25.71    | 4.10     |
| 4 | 2.02 | 1.80| 1.60| 11.04  | 20.93  | 19.44    | 36.83    |
| 5 | 2.38 | 2.00| 1.80| 16.04  | 24.44  | 22.57    | 34.39    |
| 6 | 2.07 | 2.20| 2.00| 6.18   | 3.48   | 5.81     | 3.27     |
| 7 | 2.86 | 2.40| 2.20| 16.20  | 23.18  | 20.83    | 29.81    |
| 8 | 2.22 | 2.60| 2.40| 17.37  | 8.34   | 15.59    | 7.49     |
| 9 | 2.39 | 2.80| 2.60| 17.06  | 8.70   | 16.07    | 8.19     |
| 10| 2.79 | 3.00| 2.80| 7.36   | 0.21   | 8.93     | 0.25     |
| 11| 3.65 | 3.20| 3.00| 12.43  | 17.90  | 17.51    | 25.23    |
| 12| 3.63 | 3.40| 3.20| 6.30   | 11.81  | 7.09     | 13.30    |
| **AVERAGE:** | **12.74** | **12.18** | **15.96** | **16.29** |

Table 1. Forecasting Simulation
The use of MAPEMA can reduce the problem of using MAPE. First, MAPEMA can still work even if the value of actual data is zero as far as one of the actual data in the moving average period having non zero value. Second, using moving average as the denominator, the effect of outlier data can be reduce and make MAPEMA less sensitive to outlier data than MAPE. Third, the systematically under-forecast problem of using MAPE can be reduced as shown in the previous simulated forecasting (Table I). Finally, the interpretability as the main strength of MAPE can still be achieved using MAPEMA.

4. Conclusions
This paper presents many methods for measuring the accuracy. MAPE is the most popular accuracy measure for forecasting since it is intuitive, easy to be interpreted and can be applied to measure or compare accuracy for both individual item and across item groups. However, the MAPE has several shortcomings. This paper proposes a modified MAPE using moving average of the actual data (MAPEMA) to reduce the drawbacks of MAPE. MAPEMA can reduce the problems of using MAPE and, on the other hand, still has the interpretability strength as in MAPE. MAPEMA can be an alternative performance measurement of forecasting method especially for time series forecasting. However, more researches need to be done to analyze the application of MAPEMA in real problems.

References
[1] W. C. Hang, “Performance of Stock Market Prediction,” Public Finance Q., vol. 59, no. 4, pp. 470–491, 2014.
[2] J. S. Armstrong and F. Collopy, “Error measures for generalizing about forecasting methods: Empirical comparisons,” Int. J. Forecast., vol. 08, pp. 69–80, 1992.
[3] R. J. Hyndman and A. B. Koehler, “Another look at measures of forecast accuracy,” Int. J. Forecast., vol. 22, no. 4, pp. 679–688, Oct. 2006.
[4] J. S. Armstrong, “Principles of Forecasting: A Handbook for Researchers and Practitioners,” p. 862.
[5] R. Fildes and P. Goodwin, “Against Your Better Judgment? How Organizations Can Improve Their Use of Management Judgment in Forecasting,” Interfaces, vol. 37, no. 6, pp. 570–576, Dec. 2007.
[6] S. Kolassa and R. Martin, “Percentage Errors Can Ruin Your Day (and Rolling the Dice Shows How),” Foresight, vol. Fall, p. 8, 2011.
[7] T. Foss, E. Stensrud, B. Kitchenham, and I. Myrtevit, “A simulation study of the model evaluation criterion mmre,” IEEE Trans. Softw. Eng., vol. 29, no. 11, pp. 985–995, Nov. 2003.
[8] A. Davydenko and R. Fildes, “Measuring Forecasting Accuracy: Problems and Recommendations (by the Example of SKU-Level Judgmental Adjustments),” in Intelligent Fashion Forecasting Systems: Models and Applications, T.-M. Choi, C.-L. Hui, and Y. Yu, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 43–70.
[9] S. Makridakis, “Accuracy measures: theoretical and practical concerns,” Int. J. Forecast., vol. 9, no. 4, pp. 527–529, Dec. 1993.
[10] B. Billah, M. L. King, R. D. Snyder, and A. B. Koehler, “Exponential smoothing model selection for forecasting,” Int. J. Forecast., vol. 22, no. 2, pp. 239–247, Apr. 2006.
[11] P. Goodwin and R. Lawton, “On the asymmetry of the symmetric MAPE,” Int. J. Forecast., vol. 15, no. 4, pp. 405–408, Oct. 1999.
[12] P. H. Franses, “A note on the Mean Absolute Scaled Error,” Int. J. Forecast., vol. 32, no. 1, pp. 20–22, Jan. 2016.
[13] F. X. Diebold and R. S. Mariano, “Comparing predictive accuracy,” J. Bus. Econ. Stat., vol. 20, no. 1, pp. 134–144, Jan. 2002.