Measuring Time Series Forecasting Errors

Renato R. Maaliw III, *DIT*
Professor 1, College of Engineering
Southern Luzon State University
Lucban, Quezon, Philippines

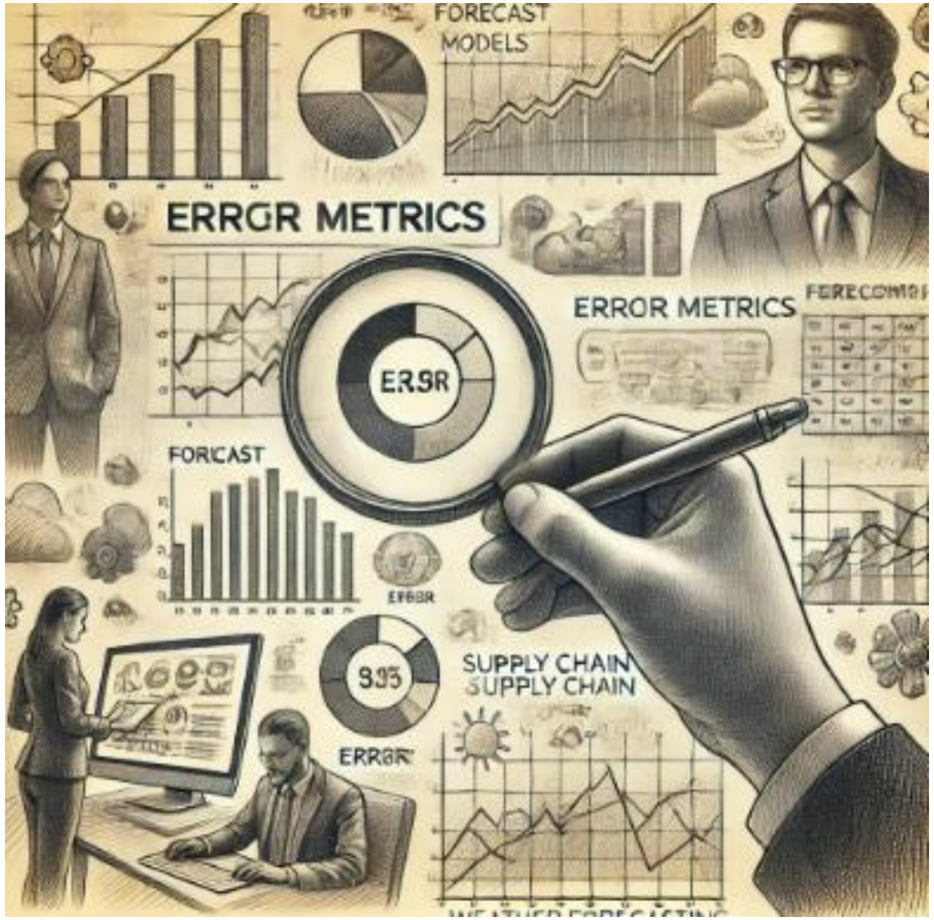
Motivation:

Forecasting is at the heart of predictions. Accurately forecasting the future drives strategic actions, yet no forecast is perfect.

Motivation:

Understanding & measuring forecasting errors is essential, it provides a quantitative way to evaluate how close our predictions to reality.

Why Measure Forecasting Errors?



1. Model Selection

Error metrics act as benchmarks that allow us to compare different model’s performance objectively.

2. Risk Management

Forecast errors are a reality, but by quantifying them, we can better assess and manage the risks involved.

Key Measures of Forecasting Errors

Mean Absolute Error (MAE)

- It gives a straightforward average of absolute differences between the predicted and actual values.

- MAE tells us, on average, how far off our predictions are in the same units as the data, which makes it easy to interpret

- $MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$

Mean Squared Error (MSE)

- It measure the average of the squares of the errors, thereby penalizing larger errors more severely.

- While MSE is effective at highlighting large errors, its squared units can make interpretation less direct.
- $MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$

Mean Absolute Percentage Error (MAPE)

- It expresses the forecast error as a percentage of the actual values, making it useful for comparing accuracy across different scales.
- It is intuitive for understanding the error in relative terms but can be sensitive when actual values are very close to zero.
- $MAPE = \frac{100\%}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right|$

Root Mean Squared Error (RMSE)

- It is the square root of the MSE, which brings the error metric back to the original units of the data.

- It is particularly popular because it maintains the benefits of MSE (penalizing larger errors) while remaining interpretable.

- $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$

Examples

Month	Actual (y_i)	Forecast ($\hat{y}_i$)
1	100	110
2	150	140
3	200	210
4	250	240

Month 1: 100 – 110 = 10 (absolute)
Month 2: 150 – 140 = 10 (absolute)
Month 3: 200 – 210 = 10 (absolute)
Month 4: 250 – 240 = 10 (absolute)

$$MAE = \frac{\mathbf{10+10+10+10}}{4} = \frac{40}{4} = \mathbf{10}$$

Interpretation: On average, the forecast is off by 10 units

Month	Actual (y_i)	Forecast ($\hat{y}_i$)
1	100	110
2	150	140
3	200	210
4	250	240

Month 1: 100 – 110 = -10
Month 2: 150 – 140 = 10
Month 3: 200 – 210 = -10
Month 4: 250 – 240 = 10

$$MSE = \frac{(-10)^2+10^2+(-10)^2+10^2}{4}$$

$$MSE = \frac{\mathbf{100+100+100+100}}{4} = \frac{400}{4} = \mathbf{100}$$

Interpretation: On average, the square of the error is 100, emphasizing larger errors.

Month	Actual (y_i)	Forecast ($\hat{y}_i$)
1	100	110
2	150	140
3	200	210
4	250	240

Month 1: 100 – 110 = -10
Month 2: 150 – 140 = 10
Month 3: 200 – 210 = -10
Month 4: 250 – 240 = 10

$$RMSE = \sqrt{\frac{(-10)^2 + 10^2 + (-10)^2 + 10^2}{4}}$$

$$RMSE = \sqrt{\frac{100+100+100+100}{4}} = \frac{400}{4} = \sqrt{100} = 10$$

Interpretation: The RMSE is 10, matching the MAE

Month	Actual (y_i)	Forecast ($\hat{y}_i$)
1	100	110
2	150	140
3	200	210
4	250	240

Month 1: $\left|\frac{100-110}{100}\right| = 0.10$

Month 2: $\left|\frac{150-140}{150}\right| = 0.0667$

Month 3: $\left|\frac{200-210}{200}\right| = 0.05$

Month 4: $\left|\frac{250-240}{250}\right| = 0.04$

$$MAPE = \frac{0.10 + 0.0667 + 0.05 + 0.04}{4} = \frac{0.2567}{4} = 0.0642$$

$$MAPE = 0.0642 * 100 = 6.42\%$$

Interpretation: On average, the forecast is off by about 6.42% relative to the actual values

Takeaway

1. **MAE** gives the average error in the same units as the data.
2. **MSE** penalizes larger errors more by squaring them.
3. **RMSE** brings MSE back to the original units, offering a clear view of the error magnitude.
4. **MAPE** expresses errors as percentage, making it easier to compare across different scales.

Various Cases

Example 1:
Sensitivity to Outliers

Month	Actual (y_i)	Forecast ($\hat{y}_i$)
1	100	105
2	100	95
3	100	100
4	100	100
5	100	160

MAE: 14
MSE: 730
RMSE: 27.02
MAPE = 14%

MAE: Forecast is off by 14 units
MSE: Squaring the errors exaggerates the outlier, resulting in high MSE
RMSE: Indicates a larger average error when the large deviation is considered
MAPE: Shows a 14% average deviation, but note it doesn't reflect the squared impact of the outlier as seen in MSE and RMSE

Insight

1. **MAE** and **MAPE** suggest a moderate average error.
2. **MSE** and **RMSE** are heavily influenced by the outlier.
3. **Voting-based decision:** In this case, if you looked at MAE or MAPE, you might underestimate the risk of the extreme error. Considering all metrics helps reveal that while most forecasts are accurate, one large error skews the overall performance significantly.

Example 2:
Near-Zero Actual Values

Month	Actual (y_i)	Forecast ($\hat{y}_i$)
1	0.1	0.3
2	0.2	0.1
3	0.15	0.2
4	0.05	0.1

MAE: 0.1
MSE: 0.01375
RMSE: 0.117
MAPE = 95.83%

MAE: Forecast is off by 0.1
MSE: Squared error remains small due to the low absolute errors
RMSE: Similar to MAE in this instance
MAPE: Appears extremely high, nearly 96%, which can be misleading given the very small scale of the actual values.

Insight

1. **MAE, MSE,** and **RMSE** indicate relatively small errors in absolute terms.
2. **MAPE** is inflated due to division by numbers close to zero, potentially distorting the true performance of the model.
3. **Voting-based decision:** By considering all metrics, you see that percentage-based errors (MAPE) may not be reliable in contexts with near-zero actuals. Relying solely on MAPE would misguide model evaluation, while the combination of MAE, MSE, and RMSE gives a more balanced view.

Conclusion

1. **Outlier Sensitivity:** RMSE and MSE can be heavily influenced by extreme values, while MAE and MAPE might understate the risk.

2. **Scale Distortion:** MAPE can be misleading when actual values are near zero, whereas MAE, MSE, and RMSE maintain their scale integrity.

Solution:

1. **Aggregate the insights:** Identify if most metrics point to similar conclusion.
2. **Understand context:** Consider the nature of your data (presence of outliers, scale issues) to decide which metrics should have more influence
3. **Balance trade-offs:** No single metric captures all nuances. A combined judgment reduces the risk of over fitting your model evaluation to a single dimension.

Other Methods:

Mean Absolute Scaled Error (MASE)

- Scales the forecast errors by the in-sample error of a naïve forecasting model (often a simple random walk).
- *Scale-independence:* It allows you to compare errors across different datasets or time series.
- *Robustness:* It doesn’t blow up when actual values are close to zero.

Other Methods:

Symmetric Mean Absolute Percentage Error (sMAPE)

- It modifies the *MAPE* by using the average of the absolute values of the actual and forecasted values in the denominator, reducing sensitivity when values are small.
- *Symmetry:* It treats over-forecasts and under-forecasts more equally
- *Reduced bias:* It mitigates the extreme percentage errors that can occur in MAPE when actual values approach zero.

Other Methods:

Theil's U Statistic

- It compares the accuracy of your forecasting model to that of a naïve model (such as the "no change" model). A value of less than 1 indicates that your model outperforms the naïve benchmark.
- *Benchmarking*: It provides context by comparing your model against a simple alternative, which can be particularly insightful for time series data.

[Code Demo]

Thank you very much for listening.