

Meyers Lab Report #1

Advay Vyas

September 1, 2025

Contents

1 Forecast evaluation

I plan to read [\[2\]](#), [\[1\]](#), and then summarize the results and its relevance to our current research focus.

1.1 General ideas and practices

ML metrics are very different than forecasting metrics because time series data is much messier and the regular ways of determining model success fail to measure accurately. Forecast origin is self-explanatory and forecast horizon is the section of time that we are predicting upon. Fixed origin evaluation uses the same training data each iteration and the forecasts are computed “one-step ahead.” On the other hand, rolling origin evaluation incorporates the new data into the testing set first, and then into the training set on the next iteration. In my opinion, rolling origin seems a lot better and I think that is what we use – not sure yet though.

1.2 Evaluating epidemic forecasts in intervals

1.3 Weighted interval score metric

2 Code investigation

2.1 Flusion

2.2 Local-Level-Forecasting

GBM_US_NSSP_public_state_pct.ipynb

3 Miscellaneous

3.1 Taylor polynomials for forecasting

References

- [1] Johannes Bracher et al. “Evaluating epidemic forecasts in an interval format”. In: *PLOS Computational Biology* 17.2 (Feb. 2021), pp. 1–15. DOI: 10.1371/journal.pcbi.1008618.
- [2] Hansika Hewamalage, Klaus Ackermann, and Christoph Bergmeir. “Forecast evaluation for data scientists: common pitfalls and best practices”. In: *Data Mining and Knowledge Discovery* 37.2 (2023), pp. 788–832.