



Contents lists available at ScienceDirect

## International Journal of Forecasting

journal homepage: [www.elsevier.com/locate/ijforecast](http://www.elsevier.com/locate/ijforecast)

## Discussion

## M4 competition: What's next?

Dilek Önkal

Newcastle Business School, Northumbria University, Newcastle upon Tyne, NE1 8ST, UK



Forecasting competitions have played a significant (albeit controversial) role in forecasting history (Hyndman, 2018). The series of M-Competitions started with M1, which used 111 time series (Makridakis & Hibon, 1979) and we now have the M4, which used 100,000 time series. The M4 Competition took the competitions to a different level by not only increasing the number of series and incorporating machine learning methods, but also acknowledging the critical challenges of forecast uncertainty and reproducibility. These neoteric extensions successfully serve the M4's stated objectives of (i) learning how to improve the forecasting accuracy and (ii) benefiting academics and practitioners (Makridakis, Spiliotis, & Assimakopoulos, 2018a, 2018b). In addition to applauding the M4 Competition's innovative dimensions and cardinal goals for both forecasters and users of forecasts, this commentary will suggest potential extensions for future M-Competitions.

Explicitly recognizing forecast uncertainty, the M4 asked for and evaluated prediction intervals as well as point forecasts. The communication of the uncertainties embedded in forecasts to the users of such forecasts is critical, and prediction intervals achieve this to a certain extent. A more detailed profile of such uncertainty would be provided by forecast distributions, and I hope that future competitions will take up this challenge and include comparative evaluations of the forecast distribution accuracy.

The added emphasis on reproducibility is another commendable novelty of this competition. This feature sets a new benchmark for future competitions and counteracts any criticisms regarding the plausibility of the findings by providing an open platform that enables replications.

Also, given the prevalence of big data and the digital torrents faced by many sectors, other promising directions for future competitions would be the inclusion of

both high-frequency data and considerably longer time series. These would offer new sets of challenges, but may be especially valuable in forecasting performance comparisons of hybrid and combination methods.

My first exposure to forecasting was through Makridakis and Wheelwright's 1977 visionary book. I remember how intrigued I was by their emphasis on the fundamental role of judgment in forecasting and on how judgment is infused throughout the steps of the forecasting process. As a researcher of behavioural dynamics in forecasting, I look forward to seeing this pivotal outlook reflected in future competitions in the form of clear and structured opportunities for including a forecaster's judgment and expertise. Noting that the M2 competition (Makridakis et al., 1993) did investigate this aspect of forecasting to some degree (with participants having the option to seek additional information in order to incorporate their judgment into budget and macroeconomic forecasts across 29 series), I believe that it would be very timely to return to this emphasis and include a systematized judgmental competition cutting across multiple domains.

The current competition is confined to time series with anonymised variables, thus concealing any contextual information and allowing judgment to be incorporated only through method selection. While the selection of a particular technique (or combination of techniques) constitutes one of the critical stages of the use of judgment in the prediction process, other platforms for including judgmental forecasts and judgmental adjustments in model forecasts directly would prove useful for enhancing our understanding of how to improve forecasting performances. This would also enable learning/feedback effects to be investigated via the provision of outcome and performance feedback to competitors. Examining how such 'reality checks' actually inform expressions of uncertainty (as reflected in probabilistic forecasts, prediction intervals and forecast distributions) would contribute to the competitions' goals of learning what lies behind the forecasts.

E-mail address: [dilekon@gmail.com](mailto:dilekon@gmail.com).

At a time when forecasting is facing heavy criticism and trust in forecasts appears to be dubious (Goodwin, 2017; Önkal, 2016), Spyros Makridakis continues to empower and inspire us with his innovative portfolio of M-Competitions, as we look forward to many more.

## References

- Goodwin, P. (2017). *Forewarned: A sceptic's guide to prediction*. London: Biteback Publications.
- Hyndman, R. (2018). A brief history of time series forecasting competitions. Hyndsight blog, <https://robjhyndman.com/hyndsight/forecasting-competitions/>, (last accessed 10.02.19).
- Makridakis, S., Chatfield, C., Hibon, M., Lawrence, M., Mills, T., Ord, K., et al. (1993). The M2-competition: A real-time judgmentally based forecasting study. *International Journal of Forecasting*, 9(1), 5–22.
- Makridakis, S., & Hibon, M. (1979). Accuracy of forecasting: An empirical investigation. *Journal of Royal Statistical Society Series A*, 142(2), 97–145.
- Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018a). The M4 competition: Results, findings, conclusion and way forward. *International Journal of Forecasting*, 34(4), 802–808.
- Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018b). Statistical and machine learning forecasting methods: concerns and ways forward. *PLoS One*, 13(3), e0194889.
- Makridakis, S., & Wheelwright, S. C. (1977). *Forecasting: Methods and applications*. Santa Barbara: Wiley-Hamilton.
- Önkal, D. (2016). Trust in forecasting. In *Paper presented at the 36th International Symposium on Forecasting*. Santander, Spain.

**Dilek Önkal** is Professor of Information Systems and Analytics at the Newcastle Business School, Northumbria University. Prof. Onkal holds a Ph.D. from University of Minnesota. Her research focuses on behavioral dynamics of forecasting, trust in forecasts, judgment and decision making, risk communication and risk perception with particular emphasis on multi-disciplinary interactions. She is Co-Editor of the International Journal of Forecasting, and her work has appeared in journals such as European Journal of Operational Research, Risk Analysis, Organizational Behavior and Human Decision Processes, Decision Sciences, International Journal of Forecasting, Technological Forecasting & Social Change, Judgment and Decision Making, Journal of Behavioral Decision Making, Journal of Business Ethics, Omega: International Journal of Management Science, Journal of Forecasting, and Decision Support Systems.