Integrating R and Excel for automatic business forecasting

Giovanni Millo[1,*], **Fabrizio Ortolani**[1]

1. Assicurazioni Generali S.p.A., R&D Department

*Contact author: Giovanni_Millo@generali.com

Keywords: RExcel, Forecasting.

We present a simple exercise in bridging the gap between statistics and everyday business practice, based on two powerful tools already available in the *R* system: the **forecast** package Hyndman (2011) for automatic time series forecasting and the *RExcel* add-in for MS Excel Baier and Neuwirth (2007), allowing to embed *R* functionality into spreadsheets and to interact with their built-in macro language. The application we developed makes forecasting practice accessible to those who are not familiar with statistical programs and, possibly, do not even have a sound statistical background.

Many processes inside the firm involve forecasting. Some build on models and relationships between balance sheet items, but sometimes an a-theoretical extrapolation of past tendencies is needed. As few firms can afford to have trained statisticians dedicated to supply-chain forecasting and the like, budgeting and other activities are often based on simple, heuristic extrapolation of past data. It is commonplace, especially in small enterprises, to "pick last year/month's value", either in terms of stocks or of increments, as the best estimate for the coming period.

Fully automatic forecasting of time series, based on model fitting and model comparing algorhythms selecting the 'best' model for the data at hand, provides a statistically well founded solution to the forecasting problem and can be of great use to the firm in obtaining accurate predictions for variables like sales, commodities' input needs and the like, where forecast errors cost money.

Such fully automatic procedures are implemented in a variety of commercial software. We show how an open-source solution is also very easy to set up.

The ideal way is thus to have the forecaster dealing only with Excel for data input, commmand issuing and results' retrieval, while a 'real' statistical engine transparently does the computing in the background.

Now the forecaster just has to select the data vector, press the trigger keys for showing up the userform, select the data frequency and press OK. He will get the forecasts at the end of the original series.

References

Baier, T. and E. Neuwirth (2007). Excel :: Com :: R. *Computational Statistics* *22*(1), 91–101.

Hyndman, R. J. (2011). *forecast: Forecasting functions for time series*. R package version 2.13.