

Editorial

Reproducibility

The importance of reproducibility has long been recognised in the natural sciences. However, in recent times there has been an acknowledgement that the highest standards of reliability and reproducibility are not always attained, even in the most highly respected journals. For example, *Nature* has published a number of articles highlighting such shortcomings and an announcement (Anon, 2013) promised that the *Nature* research journals would introduce editorial measures to address the issue. These measures included ensuring that key methodological details are reported and encouraging authors to be more transparent, for example by including their raw data.

In management mathematics, similar concerns have been expressed about the lack of research that is fully reproducible, in the sense that independent researchers are able to obtain the same results as the original study (using the same data and the same methods). For example, Boylan *et al.* (2015) pointed to the paucity of fully reproducible papers on forecasting, thereby precluding replication of findings on different data. Kendall *et al.* (2015) suggested that minimum standards are required in optimization research, to “enhance the reproducibility of results, the comparison of results, and the efficiency of individual researchers and teams”.

In response to these concerns, the editors of the *IMA Journal of Management Mathematics* asked me to work with them to develop guidelines on reproducibility for the journal. The following bullet-points summarise the agreed guidelines in terms of the conditions to be met for a publication to be recognised as “reproducible” by the award of an “R” diamond-mark:

- The first of the conditions for reproducibility is that the data used in the analysis in the manuscript must be accessible to other researchers; the detailed conditions for accessibility are discussed below.
- The second condition is that the algorithms or methods of analysis used in the manuscript should be specified in the manuscript in sufficient detail to allow reproduction of the results. The specification should be sufficiently user-friendly to allow rapid comprehension. This may be achieved using flow-charts or pseudo-code in the manuscript (or other suitable method of specification). It should not be a requirement that source code be provided. If it is provided, the more ‘user-friendly’ documentation should still be included, unless the code is so simple that it may be included in the manuscript and rapidly understood.
- The third condition is that all evaluation metrics used in the manuscript should be fully specified in the manuscript.

Judgement about the accessibility of the data used in the analysis in a manuscript are not straightforward. However, the following considerations should apply:

- For analysis based on data from a random number generator, the exact details of the generation process, and the name of the random number generator software (and version number) should be specified.
- If the dataset is available by paid subscription, then full details of the data used in the analyses in the manuscript are required, with an assurance that the data are not transitory in nature. The details should include: name of dataset, vendor, date accessed for the purpose of the calculations carried out in the manuscript and an exact specification of those elements of the dataset used.
- If the dataset is curated by a professional body or regulatory authority or such like and the journal cannot publish the data, then the manuscript should specify those elements of the dataset used and when they were accessed, and these elements should not be transitory.
- For analyses of all other data, the data should be provided in a suitable format, and with appropriate documentation. If the dataset is small, it should be included in the paper. If the dataset is too large for publication in the journal, it should be made available for publication on the journal's website.
- Author-curated datasets that are very large and that cannot be made available on the journal's website because of their size are not accessible, that is, they do not meet the conditions for accessibility.
- If any of the analyses in the manuscript refer not to the whole dataset but to some subset of it, full details of the specifications of the subsets should be provided in the manuscript.

Authors who wish their work to be recognised as “reproducible” should inform the editor in their submission cover letter. During the review process, the editors will determine what part of the manuscript will be required to be published as an electronic companion to the final article in order that the conditions for reproducibility can be met within the print constraints of the journal. The electronic companion will not appear in the print version of the journal, but will instead be published online with the accepted article.

It is hoped that the introduction of the new “R” diamond-mark will encourage the submission and publication of more fully reproducible research articles in the journal. This should contribute to lively debates about research results, and contribute towards the strengthening of our discipline.

John E Boylan

Department of Management Science
Lancaster University
Lancaster
LA1 4YX

j.boylan@lancaster.ac.uk

References

Anon (2013) Announcement: reducing our irreproducibility. *Nature*, **496**, 398.

Boylan JE, Goodwin P, Mohammadipour M and Syntetos AA (2015) Reproducibility in forecasting research. *Int J Forecast.*, **31**, 79-90.

Kendall G, Bai R, Błazewicz J, De Causmaecker P, Gendreau M, John R, Li J, McCollum B, Pesch E, Qu R, Sabar N, Vanden Berghe G and Yee A (2015). Good laboratory practice for optimization research. *J Opl Res Soc*, doi:10.1057/jors.2015.77.