A Comparison of Three Prominent Journal Metrics with Expert Judgement of Journal Quality

Peter Haddawy* and Saeed-Ul Hassan**

* *peter.had@mahidol.ac.th*
Faculty of ICT, Mahidol University, Phuttamonthon 4 Road, Nakhon Pathom, 73170 (Thailand)

** *saeed-ul-hassan@itu.edu.pk*
Scientometrics Lab, Information Technology University - Punjab, Ferozepur Road, Lahore, 54770 (Pakistan)

Introduction

In response to the demand for accurate measures of journal impact, quality, and prestige, numerous refinements of the traditional Journal Impact Factor (JIF) have been developed. Two prominent alternatives are the Source Normalized Impact per Paper (SNIP) and the SCImago Journal Rank (SJR) (Colledge et. al., 2010). SNIP is similar to JIF but corrects for differences in topicality between subject fields (Moed, 2010). It is a ratio of a journal's citation impact and the citation potential of its subject field. A journal's subject field is defined as the collection of articles citing the journal. SJR, inspired by Google's PageRank algorithm, is intended as a measure of a journal's prestige. It recursively assigns higher weight to citations from journals that are highly cited (González-Pereira et. al., 2010). Both SNIP and SJR use citation windows of 3 years, while JIF uses a citation window of 2 years.

Arguments for the appropriateness of SNIP and SJR have been made based on the logic underlying their design and studies have been carried out comparing statistical properties of SJR, SNIP, and JIF (Colledge et. al., 2010). But if these metrics are to be used as measures of journal quality, then it is also important to assess the extent to which they agree with human perception of quality. While small scale discipline-specific studies comparing JIF with expert judgement of journal quality have been carried out (Rousseau, 2008), no extensive multi discipline study has yet been carried out comparing alternative journal metrics with expert judgment. Such a study requires a sizable database of journals spanning a broad array of fields, rated by experts in the various fields. Precisely such a rating exercise was carried out by the Australian Research Council as part of its 2010 Excellence in Research for Australia (ERA) initiative. In that exercise journals were assigned to four tiers A*, A, B, C based on the perceived quality of their papers[1]. The process of producing the ranked list of 20,712 journals began in 2007 with a ranking exercise by four Learned Academies and a number of discipline peak bodies and was finalized in the consultation phase in 2010 that involved over 700 expert reviewers[2]. In this paper we study the correlation between the ERA rating and the quantitative journal metrics SJR, SNIP & JIF.

[1] http://www.arc.gov.au/era/tiers_ranking.htm

[2] The use of the ranked journal list was removed from the ERA exercise in 2011 not due to problems with the quality of the exercise but rather because "there is clear and consistent evidence that the rankings were being deployed inappropriately within some quarters of the sector, in ways that could produce harmful outcomes, and based on poor understanding of the actual role of the rankings." Senator Kim Carr, Minister for Innovation, Industry, Science and Research, May 2011.

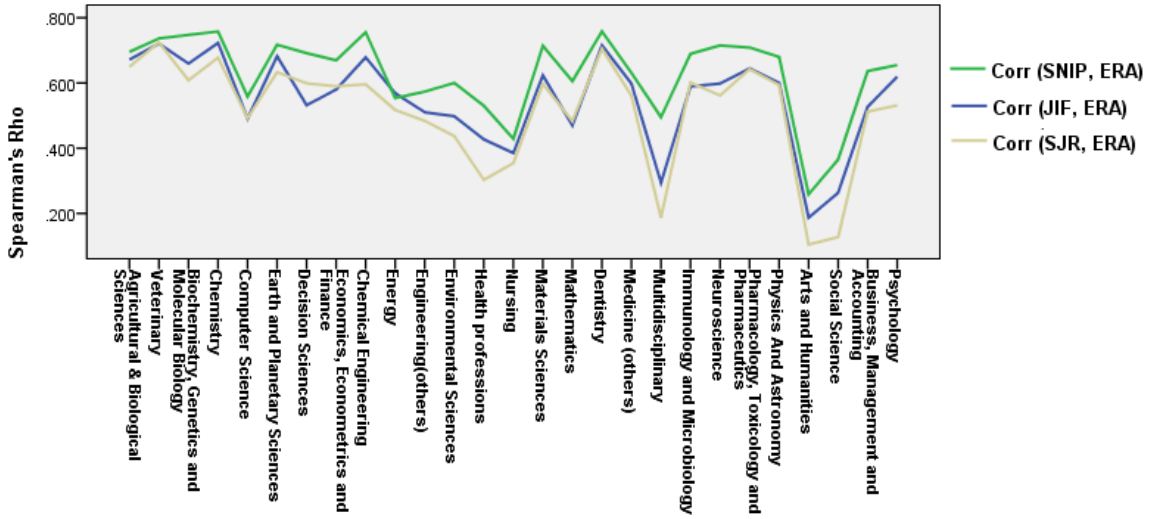
Data Collection & Methods

The 2010 SNIP and SJR metrics were downloaded from www.journalmetrics.com (retrieved 12 January 2013). We computed JIF by applying the definition to the Scopus database. Since SNIP and SJR are defined over Scopus, this controlled for the effect of the database in the comparison of the metrics. JIF for 2010 was computed by taking the ratio of the number of citations in 2010 to citable items in 2008 and 2009 divided by the number of citable items. Citable items are taken to be articles, reviews, proceedings, and notes. We identified those journals in the 2010 ERA list that are indexed in Scopus to produce the list of 11,137 journals for this analysis. We utilized the All Science Journal Classification (ASJC) to group journals for analysis by subject area. We analysed the correlation of JIF, SJR, and SNIP with the ERA rating using the Spearman's coefficient (ρ) overall and in each of 27 subject areas. We used SPSS v. 2.1 to compute the statistics.

Results and Discussion

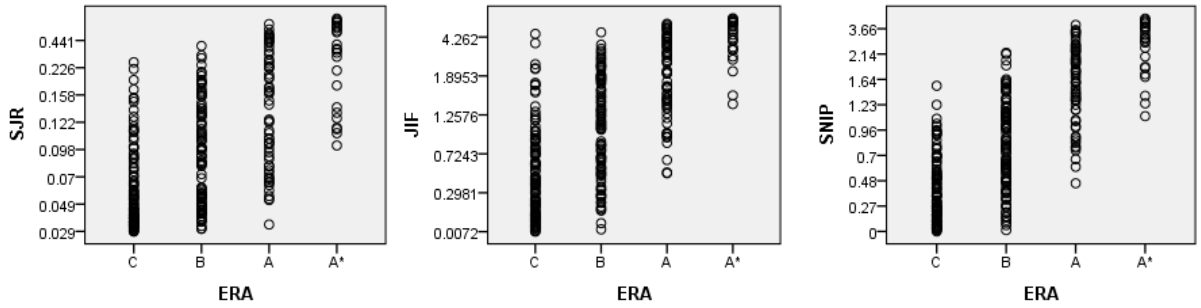
Among the selected metrics, SNIP shows the highest correlation with the ERA rating (ρ = *.537*), followed by JIF (ρ = *.374*) and then SJR (ρ = *.222*). The results are statistically significant at the .000 level, with N=11,137. Figure 1 shows the correlations of the three metrics with the ERA rating broken down by subject area. In every subject area except Energy SNIP has higher correlation than the other two metrics. SNIP has highest correlation in the areas of Dentistry (ρ = *0.758*), Chemistry (ρ = *0.758*), and Chemical Engineering (ρ=0.755). Not surprisingly, the correlation of all three metrics is lowest in the areas of Arts and Humanities, Social Science, and Multidisciplinary.

Figure 1: Correlation by subject area



More insight can be gained by viewing scatter plots of the metrics against the ERA rating. Figure 2 plots the journal metric values of 280 journals indexed under the area Chemical Engineering against the ERA rating. The correlations of the three are relatively high yet differ significantly as well: SNIP (ρ=0.755), JIF (ρ=. 678), SJR (ρ=. 595). All three metrics seem to do a better job at differentiating between A*, A, and B than between B and C while SNIP is the only metric that shows no overlap in values of A* and C journals.

Figure 2: Scatterplots of the three metrics versus the ERA rating in Chemical Engineering



Concluding Remarks

Among the three metrics, SNIP has the highest correlation with the ERA rating, followed by JIF and then SJR. This is despite the fact that one might expect the judgements of the experts to be influenced by their knowledge of the impact factors of journals. The dominance of the correlation with SNIP may have to do with the fact that the ERA rating is focused on journal quality rather than popularity so that journals could be rated highly even if they are in subfields with low citation rates. SNIP is the only one of the three metrics that normalizes for differences in citation potential across fields and subfields.

Acknowledgement

We thank Henk Moed for his guidance in mapping ERA ranked outlets to the ASJC classification.

References

Colledge, L., Moya-Anegón F., Guerrero-Bote, V., López-Illescas, C., Aisati, M. & Moed, H. (2010). SJR and SNIP: two new journal metrics in Elsevier's Scopus. *Serials: The Journal for the Serials Community*, 23(3), 215-221.

González-Pereira, B., Guerrero-Bote,V., Moya-Anegón, F. (2010). A new approach to the metric of journal's scientific prestige: The SJR indicator. *Journal of Informetrics*, 4(3), 379–391.

Moed, H.F. (2010). Measuring contextual citation impact of scientific journals. *Journal of Informetrics*, 4(3), 265–277.

Rousseau, S. (2008). Journal evaluation by environmental and resource economists: A survey. *Scientometrics*, 77(2), 223–233.