Earthscientometrics for Editors and Authors: Digital Dreams or Dilemmas?

– K.R.Y. Simha, Indian Institute of Science (IISc), Bengaluru - 560 012 (*E: simha@iisc.ac.in*)

Scientometric appraisal of scientific publishing has become as much a source of excitement as resentment among authors, editors, publishers and readers alike. Achieving global recognition with local resources is the stuff of legendary history in sports, medicine, science and mathematics. Editors and authors alike are striving for showcasing papers on a global scale while publishers are exploring new technologies and business strategies to promote their journals. It is against this backdrop that the *JGSI* analysis searching the Web of Science by Renjit and Pradeepkumar in February 2021 DigitalCommons@UN-Linclon (hereafter RP) makes interesting reading deserving a larger discussion. The RP data search for the period 1989-2016 lists top 100 papers by 272 authors cited 32 to 165 times for a grand total of 5637 (average authorship n=2.72); and, this note offers further interpretation and appreciation of Earthscientometrics (ESM).

The ESM score card unveils the synergistic editor-reader-author troika attracting the national, international and global journal readership. The readership for national journals published by societies is about twice the active membership (about 5,000 for *JGSI*). This increases tenfold for international journals and hundred fold for multidisciplinary global brands like Nature. Therefore, enlarged samples (1000 for international and 10,000 for global) are required to interpret and reveal trends for comparative ESM at tempted by RP.

Sample size and scaling as the key aspects of statistics and scientometrics inspired the h-index concept. Introduced in 2005, the h-index has captured the imagination of computer scientists and statisticians alike for exploiting Artificial Intelligence/Machine Learning (AI/ML) decision tools for ranking individuals, institutions and journals. The h-index defined simply as the number N of papers of attracting N or more citations can also be applied to a journal assuming the editor wrote them all! The RP scatter diagram suggests 47 papers with more than 47 citations giving the *JGSI* h index. Relevant ESM data for training AI/ML codes include: total citations (C), paper total (T), mean (m= C/T), peak citation (p) and author total (A) enable extending RP through correlating (h, C, T, m, p, A).

Firstly, comparing *JGSI* with an international journal like Annals of Surgery (AoS), scaling up tenfold gives an average of 563.7 tallying well with AoS average of 609. Conversely, this predicts that AoS average of the top 1000 papers scales down to 60.9 consistent with RP.

Secondly, the Journal of Neuro Interventional Surgery (JNIS) cited by RP with a median of 26 citations is basically a national journal with 79% articles originating in the USA. Paradoxically on the other hand, top 100 papers in the journal of Pediatric Surgery International (PSI) mentioned by RP drew only 3309 citations. This may be due to its shorter existence since 1989; and, perhaps PSI will draw level with others after 40 years. As a corollary, scaling up tenfold also implies that effectively *JGSI* h index rises to 470 if deemed as in international journal with affiliations to international societies and agencies.

Thirdly, GSI and NGRI leading the RP list highlight practitioners

contributing to the growth of *JGSI* as a premier journal attracting academicians in India and abroad. Academicians do enjoy higher citations especially in popular areas of chemistry and computer science (CS), but mind boggling citations for Lowry (C>260,000) and Perdew (C>134,000) without a Nobel Prize defies logic! Among 2 million CS authors in a 2018 survey, only 1000 (0.05%) had h>40 and just 87 with h>87. In this context, recognizing Stanford top 2% and labelling h>100 papers as citation classics by RP seem forty-fold lax and imprudent. The anomalous Type II error in the case of Santosh raised by RP is a clarion call for blending ESM with AI/ML to avoid unpleasant shocks.

Fourthly, that 272 authors wrote the top 100 papers vehemently demonstrate small authorships averaging 2.7 guided by vigilant editors can generate h=47 – a level attained with much larger authorships. Assigning a higher weightage for smaller authorships seems prudent for promoting original papers by students and young scientists. Higher weightage for single and first authorship can be implemented using a frequency and byline (b) position index h(A, b). The RP methodology could help model h(A,b) for advanced ESM research through visualization of areas of expertise and other academic affiliations.

Fifthly, as a rapid AI/ML primer, imagine C obeys a power law $C = h^s$. Then, s determined from (h,C) data is about 2.4. A cursory scan of websites reveals C>100 correlating with h = 7-8 and C>500 with h = 13-15. The milestone C = 1,000 correlates with h = 16-18. Only a few succeed C > 5,000 carrying h>35. Only beyond T>300, it is possible to expect (h/C) > (45/10K); and, (h/T/C) > (90/600/50K) is intense global networking publishing 30 papers annually in popular subjects. Amazingly, $C = h^{2.4}$ correlates well beyond (h, C) > (100,100K) save for a few freaks a la Lowry and Perdew!

Sixthly, extending AI/ML to ESM for correlating (h, m, p), again a power law $mp = h^w$ lends support with w = 1.8 for a select range of data culled from websites. These power law exponents w and s have to be periodically tweaked and calibrated to suit different databases like WoS, Google Scholar etc. Further, AI/ML codes warrant periodical updating and fire walling against hacking and ransomware afflicting the digital world.

Seventhly, outstanding papers in national journals remain uncited eclipsed by global papers periodically reporting on climate and high energy physics including a ludicrous 5,154 author paper in Physical Review Letters! Defining a citation reliability quotient $q = h/\sqrt{C}$ can help resolve Type II errors with q < 0.2 hinting spurious authorship.

Finally, 10% annual membership growth can propel *JGSI* towards international levels of h, C, T, m, p, q, s, w etc. As suggested by RP, inviting overseas authors with Indian education or work experience can attract more citations. Ramping up regular submissions showcasing private and public R&D missions driving the Indian economy can add value to *JGSI*. Incentives for doctoral candidates enabling publishing their work deserve special attention. Overall, *JGSI* editors have managed admirably over the decades navigating the dizzy sea of digital technology while retaining the core values of rigour and regularity. Kudos to JGSI, ESM, AI/ML and digital publishing!