BRIEF COMMUNICATION

Some Comments on "A Proposal for a Dynamic h-Type Index" by Rousseau and Ye

Quentin L. Burrell
Isle of Man International Business School, The Nunnery, Old Castletown Road, Douglas, Isle of Man IM2 1QB, via United Kingdom. E-mail: q.burrell@ibs.ac.im

Caution is urged over the adoption of dynamic h-type indexes as advocated by Rousseau and Ye (2008). It is shown that the dynamics are critically dependent upon model assumptions and that practical interpretation might therefore be problematic. However, interesting questions regarding the interrelations between various h-type indexes are raised.

Introduction

In their recent paper, Rousseau and Ye (2008; referred to as R&Y henceforth) seek to make use of the way that various h-type indexes develop over time, perhaps to differentiate between scientists whose impact is increasing rather than staying constant or decreasing. The way they suggest doing this is to assume that the indexes are continuous functions of time and to consider the derivative to measure rate of change. This is fine except that the h-index, for instance, does not increase (as a function of time) according to any continuous function since it can only have integer increments—a form commonly referred to as a step function. Similar reservations hold for the other h-type indexes. Hence the best one can hope for is a reasonable continuous approximation.

The Background

In the paper in which he first formulated the h-index, Hirsch (2005) gave a heuristic argument suggesting that, for a particular (publishing) author, the index would increase linearly over (discrete) time. Burrell (2007a), extending a previously published stochastic model (Burrell, 1992) for an author's production/citation process, indicated, on the basis of numerical investigations, that the (expected) index should indeed increase (approximately) linearly with time under a wide set of variations in the publication and citation rates. Some empirical support for the assumption of a linear form was provided by Burrell's (2007b) analysis of Liang's (2006) data, although there were cases where the linear form clearly did not seem appropriate and for which possible scientometric explanations were proposed. Hirsch (2007) also provided examples of the approximate linear development.

The Rousseau and Ye Study

In Table 1 of R&Y are given the values of Rousseau's own h-index (Hirsch, 2005), rational h-index (hrat; Ruane & Tol, 2008), and R-index (Jin et al., 2007) over the period 2001–2008. The growths of these indexes are plotted in Figure 1, with Year 0 corresponding to 2001 and so on, through to 2008 being Year 7. For all three, there seems to be no good reason to assume other than that the growth is (approximately) linear. Indeed we have included the fitted least-squares regression line constrained to pass through the origin, as it should, and in each case find an R^2 value of .97 or more.

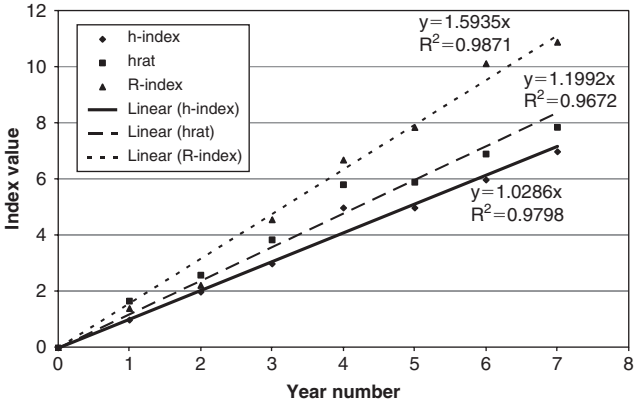

FIG. 1. Time development of Rousseau's h-type indexes.

Received July 21, 2008; revised August 27, 2008; accepted August 27, 2008

 DOI: 10.1002/asi.20969

For the rational h-index, R&Y instead propose—but without any supporting argument—fitting a power law ($y = 1.67x^{0.801}$) and find $R^2 = .984$, a slight improvement on the $R^2 = .967$ for the linear fit ($y = 1.1992x$). It should be noted, however, that the linear model through the origin is a one-parameter model while the power law has two parameters. Indeed, if we are just engaged in a curve-fitting exercise, a quadratic curve constrained to pass through the origin is another two-parameter model ($y = -0.062x^2 + 1.547x$) but with $R^2 = .988$! Hence we have three different curves giving good fits to the rational h-index. To find the corresponding dynamic h-index (as defined by R&Y) for each we need the coefficient determined by the derivative of the function at time $t = 7$. For the linear fit this is just the slope 1.20, for the power law it is 0.91, and for the quadratic it is 0.68. Hence the dynamic h-index is extremely sensitive to the choice of even the type of fitted function.

We would suggest that, although long-term changes in h-type indexes, as alluded to in R&Y and illustrated in Burrell (2007b), are important, the instantaneous change approach advocated by R&Y may be problematic in practice. As we have shown, for the sort of noncontinuous data being considered it is extremely sensitive to the underlying assumptions. Unless there are good a priori reasons for choosing a particular functional form, results should be interpreted with caution.

Further Considerations

Even though it is a case study over a fairly short period of time, R&Y's Rousseau data suggest some interesting and, we believe, possibly important questions. First, each of the indexes seems to be (approximately) proportional to time and hence each is (approximately) proportional to each of the others. It is also worth mentioning that the stochastic model suggests that Jin's (2006) A-index is also approximately proportional to the h-index, and hence possibly with the others (Burrell, 2007c). This would indicate that all of these indexes are essentially measuring the same phenomenon up to a scaling factor! In this case it would seem that for practical purposes the best index is the one that is easiest to calculate.

Second—and this is a much deeper problem—if it is indeed true that these indexes are (even approximately) proportional to each other, in what way do the constants of proportionality depend upon, for instance, the parameters of the publication and citation processes? Many more empirical studies along the lines of R&Y may well provide clues about the relationships, but a theoretical solution does not seem to be immediately apparent.

As a final point, one of the justifications R&Y plead for their criticism of the simple linear model refers to Egghe's (2007) mathematical (and purely deterministic) model for the evolution of the h-index. However, they overlook the fact that Egghe's model relates to a fixed body of work, i.e., it seeks to describe only the citation process. Hence it might, for instance, be appropriate for a scientist whose publication career has ended, but not for the application of interest here, namely to currently active scientists for whom new citation-attracting publications appear during the period of study. For such studies Burrell's (2007a) model, despite its acknowledged deficiencies, seems to be the most comprehensive available as yet.

Conclusion

Our view is that the linear form for the development of an author's h-index, as argued by Hirsch (2005, 2007) and Burrell (2007a), should in general be regarded as the "typical scenario" and that deviations from this should be investigated scientometrically, as suggested by Burrell (2007b), rather than by simply fitting a curve.

References

Burrell, Q.L. (1992). A simple model for linked informetric processes. Information Processing & Management, 28, 637–645.

Burrell, Q.L. (2007a). Hirsch's h-index: A stochastic model. Journal of Informetrics, 1, 16–25.

Burrell, Q.L. (2007b). Hirsch index or Hirsch rate? Some thoughts arising from Liang's data. Scientometrics, 73, 19–28.

Burrell, Q.L. (2007c). On the h-index, the size of the Hirsch core and Jin's A-index. Journal of Informetrics, 1, 170–177.

Egghe, L. (2007). Dynamic h-index: The Hirsch index in function of time. Journal of the American Society for Information Science and Technology, 58, 452–454.

Hirsch, J.E. (2005). An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences of the United States of America, 102, 16569–16572.

Hirsch, J.E. (2007). Does the h-index have predictive power? Proceedings of the National Academy of Sciences of the United States of America, 104, 19193–19198.

Jin, B.H. (2006). H-index: An evaluation indicator proposed by scientist. Science Focus, 1(1), 8–9.

Liang, L. (2006). H-index sequence and h-index matrix: Constructions and applications. Scientometrics, 69, 153–159.

Rousseau, R., & Ye, F.Y. (2008). A proposal for a dynamic h-type index. Journal of the American Society for Information Science and Technology, 59, 1853–1855.

Ruane, F.P, & Tol, Richard S.J. (2008). Rational (successive) h-indices: An application to economics in the Republic of Ireland. Scientometrics, 75, 395–405.

DOI: 10.1002/asi