Available online at www.sciencedirect.com
ScienceDirect
Journal of Informetrics 1 (2007) 16–25
Journal of INFORMETRICS An International Journal
www.elsevier.com/locate/joi

Hirsch's h-index: A stochastic model

Quentin L. Burrell *

Isle of Man International Business School, The Nunnery, Old Castletown Road, Douglas, Isle of Man IM2 1QB, via United Kingdom

Received 29 June 2006; received in revised form 27 July 2006; accepted 28 July 2006

Abstract

We propose a simple stochastic model for an author's production/citation process in order to investigate the recently proposed h-index for measuring an author's research output and its impact. The parametric model distinguishes between an author's publication process and the subsequent citation processes of the published papers. This allows us to investigate different scenarios such as varying the production/publication rates and citation rates as well as the researcher's career length. We are able to draw tentative results regarding the dependence of Hirsch's h-index on each of these fundamental parameters. We conjecture that the h-index is, according to this model, (approximately) linear in career length, log publication rate and log citation rate, at least for moderate citation rates.

Keywords: Hirsch h-index; Stochastic model; Informetric process

1. Introduction

Hirsch (2005) proposed the h-index, a single number to measure an individual's research output and its impact. In the original (preprint) version, his definition states (essentially) "*A scientist has index h if h of his/her papers have at least h citations each and the rest have fewer than h citations each*". Since its first publication, the index has received much attention, both in the popular domain, as in Ball (2005), and in the academic literature. In the latter we can distinguish straightforward attempts to apply the index, as in Bornmann and Daniel (2005), Rousseau (2006) and Cronin and Meho (2006); those seeking to modify the index or extend its range of application, as in Braun, Glänzel, and Schubert, 2005, and Egghe (2006); those that relate the h-index to the "traditional" bibliometric evaluation measures, such as Glänzel (2006b) and Van Raan (2006), and those seeking to give some sort of mathematical model for the index, as in Egghe (in press), Egghe and Rousseau (in press), Glänzel (2006a), as well as the original paper of Hirsch (2005). In this paper we use an established informetric way of modelling the production/citation process to seek to give some insight to the index.

2. The publication-citation model

Consider an author whose publishing career begins at time zero and we then wish to model the numbers of citations to each of his/her publications by time T (the present). (Already, we have a small problem as to how we define "time

* Tel.: +44 1624693706; fax: +44 1624665095.
E-mail address: q.burrell@ibs.ac.im.

1751-1577/$ – see front matter
doi:10.1016/j.joi.2006.07.001

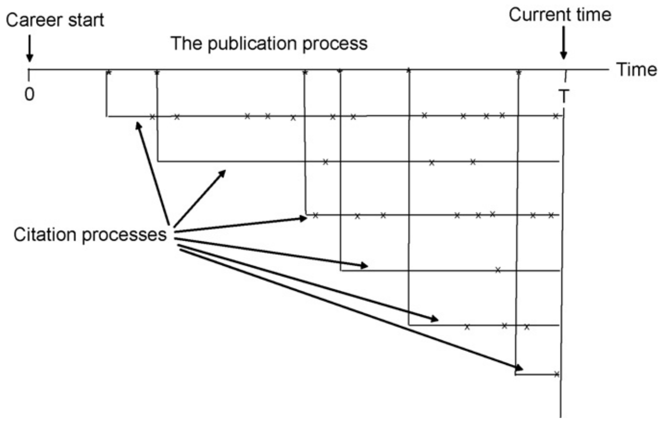

Fig. 1. Representation of the publication/citation processes.

zero". Is it the time at which the author takes up his/her post, or when the first publication is submitted, or accepted, or appears in print? In what follows, let us not worry over this detail.) In other words we are assuming that currently the individual is T time units into his/her productive career. (The unit of time will typically be 1 year, but for modelling purposes this is not important.) Thus, we assume that the author publishes papers at certain times and that these papers subsequently attract citations following their publication, where both the publication and citation accumulation processes are random. We further assume that some papers are more citable than others so that the citation rate varies between different publications. A schematic representation of the model is given in Fig. 1 where we have an author with six papers published during the period of observation. The first (earliest) of these gains 12 citations, the second 3, then 8, 1, 3 and 1 for the rest.

In order to make this general scenario analytically viable we need to be more precise, so our initial model assumes:

Assumption 1. From the start of his/her publishing career at time zero, an author publishes papers according to a Poisson process of rate θ. Thus, by time T, the number of publications Y_{T} has the distribution

$$P(Y_{\mathrm{T}}=r)=\mathrm{e}^{-\theta T}\frac{(\theta T)^r}{r!},\quad r=0,1,2,\ldots,\quad \text{and}\quad E[Y_{\mathrm{T}}]=\theta T$$

Note that the parameter θ gives the mean number of publications per unit time for this author, called the *publication rate*.

Assumption 2. Any particular publication acquires citations according to a Poisson process of rate Λ, where Λ varies from paper to paper. Here, Λ denotes the mean number of citations per unit time following publication, called the *citation rate*.

Assumption 3. The citation rate Λ for this author varies over the set of his/her publications according to a gamma distribution of index $\nu \geq 1$ and scale parameter $\alpha > 0$. Thus, the probability density function of Λ is given by

$$f_{\Lambda}(\lambda)=\frac{\alpha^{\nu}}{\Gamma(\nu)}\lambda^{\nu-1}\mathrm{e}^{-\alpha\lambda},\quad 0<\lambda<\infty$$

Note that $E[\Lambda] = \nu/\alpha$ gives the overall *mean citation rate*, or the average number of citations acquired by a randomly selected paper of this author per unit time.

Remarks.

(i) The usual requirement for the gamma index ν is that it is positive. We require the restriction $\nu \geq 1$ in order to ensure convergence of certain integral expressions in the following. As this should be viewed as an exploratory model, the restriction is of no real concern.
(ii) This model has previously been suggested by Burrell (1992) to develop an idea of Rousseau (1992) and can be considered as an example of what Egghe and Rousseau (1990, p. 378) call a three-dimensional informetric study.

In Burrell (1992) the focus of interest was simply the total number of citations accumulated by the collection of publications, here we are interested in the distribution of the number of citations to the individual publications.

For any author, the "current time" T is equivalent to the time since the author's publication career began, which we take as defining time zero (for this author). For our author, let us therefore denote by X_T the number of citations achieved by a (randomly chosen) published paper by time T and by $N(n; T)$ the number of published papers receiving at least n citations by time T. Then we have:

Theorem 1. *Under the assumptions of the model, the distribution of the number of citations to a randomly chosen paper by time T is given by*

$$P(X_T = r) = \frac{\alpha}{(\nu - 1)T} B\left(\frac{T}{\alpha + T}; r+1, \nu - 1\right) \text{ for } r = 0, 1, 2 \ldots \quad (1)$$

where

$$B(x; a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^x y^{a-1}(1-y)^{b-1} \mathrm{d}y$$

is the cumulative distribution function of a beta distribution (of the first kind) with parameters a and b.

Proof. See the Appendix.

So far as determination of the h-index is concerned, our main result is almost a corollary of the above:

Theorem 2. *Under the assumptions of the model, the expected number of papers receiving at least n citations by time T is given by*

$$E[N(0;\ T)] = \theta T$$

$$E[N(n; T)] = \theta T \left(1 - \frac{\alpha}{(\nu - 1)T} \sum_{r=0}^{n-1} B\left(\frac{T}{\alpha + T}; r+1, \nu - 1\right)\right) \text{ for } n = 1, 2, 3, \ldots \quad (2)$$

Proof. The proof of this result follows from a pair of Lemmas, the proofs of which are given in the Appendix. The result for $n = 0$ is of course just the expected number of publications by time T. □

Lemma 1.

$$E[N(n; T)] = \frac{\alpha\theta}{(\upsilon - 1)} \sum_{r=n}^{\infty} B\left(\frac{T}{\alpha + T}; r+1, \nu - 1\right)$$

Lemma 2.

$$\sum_{r=0}^{\infty} B\left(\frac{T}{\alpha + T}; r+1, \upsilon - 1\right) = \frac{(\upsilon - 1)T}{\alpha}$$

Remarks.

(a) Note that for given parameter values and T, evaluation of the right hand side of (2) is straightforward with any statistical package including evaluation of the cumulative distribution function of the Beta distribution.
(b) Although we have adopted a notation emphasising the dependence on n and T, the expected number of citations actually depends on the model parameters θ, α and ν so perhaps we should write $E[(n; T)|\theta, \alpha, \nu]$. Although this is rather cumbersome, it simplifies when we see from (2) that

$$E[(n; T)|\theta, \alpha, \nu] = \theta E[(n; T)|1, \alpha, \nu] \quad (3)$$

This is useful in later calculations.

(c) The following are intuitively obvious, but note from (2) that the expected number of papers receiving at least n citations
    (i) is proportional to θ, the publication rate. Hence, all other things being equal, more productive authors achieve more (expected) citations.
    (ii) is increasing in T for any n.
    (iii) is decreasing in n for any T.

3. The h-index

The h-index as originally defined for an individual scientist is that number h such that h of his/her papers have at least h citations, while the rest have no more than h citations. Although Hirsch (2005) wrote in terms of papers written by scientists, there is no reason not to extend this to any academic discipline, hence we talk of authors. Also, Glänzel (2006a) has pointed out that there is a possible ambiguity in the case where an author has several papers with the same number of citations at h. (This was realised by Hirsch (2005) in the published version of his paper.)

Thus, the h-index is concerned with the distribution of the number of citations accumulated by each of an author's publications up to the current time. In our notation, Hirsch's index is given by:

Definition 1. Hirsch's h-index at time T is, for any particular author, the integer $h(T)$ satisfying

$$h(T) = \max\{n : n \leq N(n; T)\}$$

Note that this is the modified form proposed by Glänzel (2006a) to take account of possible ties. Also, this is an empirical measure, requiring observation of the actual values of $N(n; T)$. Let us remark at this stage that the upper bound for any author is his/her total number of publications.

Here we are considering a theoretical model so let us modify the above to:

Definition 2. The theoretical h-index at time T is the integer $h(T)$ satisfying

$$h(T) = \max\{n : n \leq E[N(n; T)]\}$$

Note that this is well defined since we have already remarked that $E[N(n; T)]$ decreases with n. Also, since $E[N(n; T)]$ is the expected value of a random variable and hence is not necessarily an integer it will be useful to define

$$h^*(T) = E[N(h(T); T)]$$

being the expected number of papers receiving at least $h(T)$ citations and note that necessarily $h^*(T) \geq h(T)$.

4. Exploring different scenarios

Our model involves four parameters:

(i) The author's publication rate, θ.
(ii) The gamma parameters, ν and α.
(iii) The length of the author's publishing career to the current time, T.

All four of these parameters can vary from author-to-author. So far as the publication rate is concerned, some authors are very prolific, others less so. For the gamma parameters, note that ν/α is the average citation rate for this author's papers. This average citation rate may reflect the author's stature in the field but may well also be dependent on the field as different fields have different citation practices. In addition, those working in smaller subject areas are addressing smaller audiences from whom citations are generated. Clearly the length of the publishing career also varies. We will consider various scenarios to illustrate the dependence on the individual parameters.

Table 1
Determination of the theoretical h-index: $E[N(n; 10)]$ for varying production rate, θ

n	$\theta = 2$	$\theta = 5$	$\theta = 10$
0	20.00	50.00	100.00
1	19.50	48.75	97.50
2	19.00	47.50	95.00
3	18.50	46.25	92.50
4	18.00	45.00	90.01
5	17.50	43.76	87.52
6	17.01	42.52	85.03
7	16.51	41.28	82.55
8	16.02	40.04	80.09
9	15.53	38.82	77.63
10	15.04	37.60	75.20
11	14.56	36.39	72.78
12	14.08	35.19	70.38
13	**13.60**	34.00	68.01
14	13.13	32.83	65.66
15	12.67	31.67	63.35
17	11.76	29.41	58.22
19	10.89	27.22	54.45
21	10.05	25.13	50.25
23	9.25	**23.12**	46.24
25	8.49	21.21	42.43
27	7.76	19.42	38.82
29	7.08	17.71	35.42
31	6.44	16.12	**32.24**
33	5.85	14.63	29.27

4.1. Varying the publication rate

To illustrate the determination of the h-index we consider first a trio of authors who differ only in their publication rate. Suppose that these three publish on average $\theta = 2$, 5, and 10 papers per year, respectively, and that the gamma parameters are $\alpha = 1$, $\nu = 5$ for all three so that each paper (for each author) receives five citations per year on average (after publication). Then with a publishing career of (current) length $T = 10$ years for each, calculations based on (2) lead to the results in Table 1. (Note that we have trimmed and truncated the output for the current purpose.)

Of course, since $E[N(n; T)]$ is directly proportional to θ, see (3), the columns are simply multiples of the one that could be calculated for $\theta = 1$. From the table we can read off the highlighted h-index (expected h-index) in the three cases as $h = 13$ ($h^* = 13.60$), $h = 23$ ($h^* = 23.12$) and $h = 31$ ($h^* = 32.24$), respectively. As should be expected on intuitive grounds these increase with the publication rate but note that the increase is nonlinear. This can be seen in Fig. 2(a) where the h-index for $T = 10$ is plotted against the production rate from 0 to 50, with the same gamma parameters as above. In fact we plot both h and h^* for illustration, confirming that always $h^* \geq h$.

Note that we have only considered values of θ up to 50, i.e. an author producing on average an article practically on a weekly basis, which was felt to be a reasonable upper value in practice. Fig. 2(b) shows the same data but with a logarithmic scale for the production rate and, for clarity, using only the h^* values. Here, we see almost perfect linearity. In fact, note that we have included productivity values θ as large as 500 and this does not cause any deviation from the same linear fit! (For the actual values plotted the R^2 value is greater than 0.99.) Further numerical investigation suggests that this result is robust for other values of both T and the gamma parameters. Hence, our model suggests that the h-index is approximately linear in the logarithm of the publication rate.

4.2. Varying the career length

By its very definition, an author's h-index cannot decrease in time so here we investigate its time dependence. We note that Egghe (2006) considers a time-dependent model for the h-index but his model is very different from ours

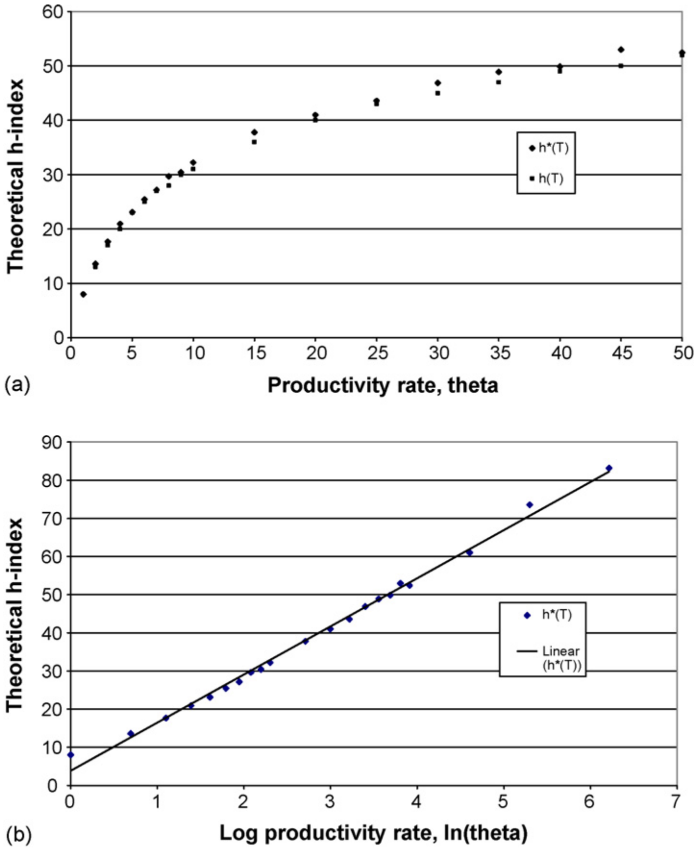

Fig. 2. (a) h-index as a function of productivity. (b) h-index as a function of log productivity.

in that he essentially supposes that an author's entire body of work is available at time zero and it is the subsequent accumulation of citations that is investigated. Similarly, Glänzel (2006a) does not model the productivity process, only the citation distribution. On the other hand, our model allows new publications to appear during the period of observation so that we are genuinely modelling the evolution of an author's publication/citation career. Given the three different author productivities and gamma parameters as in the previous section, let us consider career lengths of 5, 10 and 20 years. (Note that when we speak of career length we mean "current career length", i.e. we are thinking of an author who is currently active but whose productive career began 5, 10 or 20 years ago.) Rather than presenting the numerical analysis we adopt a graphical approach which perhaps gives greater insight into the interaction between time and productivity.

In Fig. 3 we have plotted $E[(n; T)]$ for the three different career lengths and for $\theta = 1$ in each case. Notice straight away that, as expected, $E[(n; T)]$ decreases with n for each fixed T and increases with T for any n. (Remark: Although we are looking at a function of the discrete variable n, we have plotted it as a continuous function for ease of visual interpretation.) To find the approximate h-index, at least to graphical accuracy, this is given, for any θ, by the solution of

$$n = E[N(n; T)|\theta] = \theta E[N(n; T)|1],$$

or

$$n/\theta = E[N(n; T)]$$

Hence in Fig. 3 we also plot the lines n/θ corresponding to the various θ values, allowing us to read off the (approximate) h-values. The exact values are summarised in Table 2.

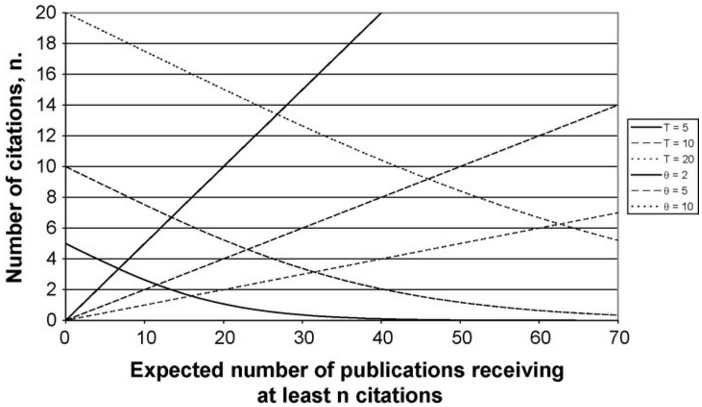

Fig. 3. h-index as a function of time and of publication productivity.

Table 2
The h-index for varying career length, T

	$T=5\ h(h^*)$	$T=10\ h(h^*)$	$T=20\ h(h^*)$
$\theta=2$	6 (7.06)	13 (13.60)	26 (27.64)
$\theta=5$	11 (12.19)	23 (23.12)	45 (46.95)
$\theta=10$	15 (17.29)	31 (32.24)	62 (63.64)

The, perhaps surprising, observation from these figures is that the h-index is (almost) directly proportional to T, the length of an author's career to the current time (given the constancy of the other parameters). Further numerical investigations suggest that this approximate result is true over a wide range of parameter values. This result was in fact conjectured by Hirsch (2005). Note that Egghe (2006), using very different model assumptions found a rather more complex time dependence.

4.3. Varying the gamma parameters

In Table 3 we give the expected h-index for production rate $\theta=5$ but different career lengths and different combinations of the gamma parameters, subject to the mean citation rate ν/α over all publications being the same. For purposes of illustration we have taken this mean citation rate to be $\nu/\alpha=10$. What is surprising here is that it is not the individual parameter values that are crucial, only their ratio so that it is the overall mean citation rate which has the greatest influence, so far as the citation process is concerned, on the h-index.

In order to investigate what the relationship between the h-index and the mean citation rate might be, consider an author with $T=20$ and $\theta=5$, so the expected total number of publications is 100. To illustrate the varying citation rate, we take $\alpha=5$ and then selected values of mean citation rate between 1 and 100 so that ν varies between 5 and 500. The results for h^* are plotted in Fig. 4(a) and, with a logarithmic scale, in Fig. 4(b).

The curve in Fig. 4(a) appears to be approaching an asymptotic value of 100. This is because, just as an author's h-index has the total number of publications as its upper bound, so the theoretical index is limited by the expected

Table 3
The h-index for varying gamma parameter values, with $\nu/\alpha=10$

	$T=5\ h(h^*)$	$T=10\ h(h^*)$	$T=20\ h(h^*)$
$\alpha=1,\ \nu=10$	23 (24.76)	47 (48.33)	95 (95.46)
$\alpha=2,\ \nu=20$	24 (24.79)	48 (49.53)	97 (97.97)
$\alpha=5,\ \nu=50$	24 (25.51)	49 (50.00)	98 (100.00)
$\alpha=10,\ \nu=100$	24 (25.76)	49 (50.51)	99 (100.00)
$\alpha=100,\ \nu=1000$	24 (25.98)	49 (50.95)	99 (100.90)

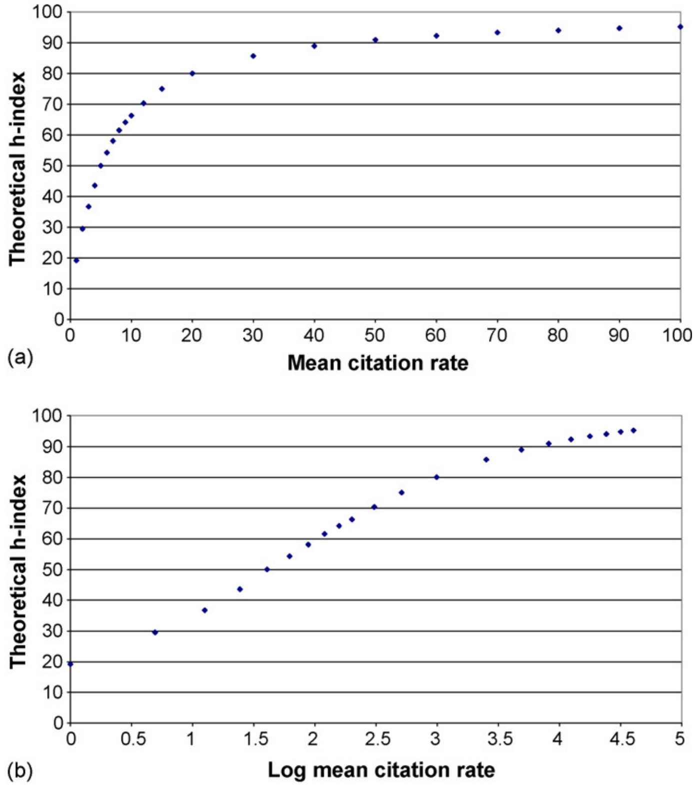

Fig. 4. (a) h-index as a function of citation rate. (b) h-index as a function of log citation rate.

number of publications, in this example 100. By the time the citation rate reaches 50, already we have $h = 90$ ($h^* = 90.96$). Turning to the log of the mean citation rate in Fig. 4(b), we have an elongated S-shape, with a levelling off at the upper end, again because of the maximal value of the h-index. However, note an approximate linearity in at least the early stages of Fig. 4(b). (For the full range of citation rates plotted, we find $R^2 = 0.970$, but clearly curvilinear. For rates up to 20, $R^2 = 0.993$ and close to linear.) Further numerical investigation suggests that, with this model, the theoretical h-index is approximately linear in the log of the citation rate, at least for “moderate” citation rates.

5. Concluding remarks

The model investigated here, based upon both the publication and citation processes being considered as Poisson processes, is perhaps the simplest genuinely stochastic model that could be considered. As such, it is worthwhile pointing out not just the general observations regarding the h-index derived from the model but also the limitations and possible modifications of the model for future work.

So far as the model is concerned, our investigations suggest that, other things being equal, Hirsch’s h-index

(i) is approximately proportional to current career length, T;
(ii) is approximately a linear function of the logarithm of the author’s productivity rate;
(iii) is approximately a linear function of the logarithm of the mean citation rate for the author, for moderate citation rates.

Obviously, because of their striking simplicity these observations are of interest, but note that they are based upon numerical and graphical investigations rather than analytical studies so that they should be viewed as conjectures and there is room for further theoretical work. And are there intuitively reasonable arguments to support any of them?

There is then much to be done to see if they do indeed reflect what actually happens for individual authors, i.e. does the model reflect reality? In particular, the role of the citation rate surely involves both the field in which the author works and the author's stature in that field—it would be interesting to investigate the interrelationship of these.

It is admitted that this is a simple model, based upon very specific assumptions all of which are open to challenge on intuitive grounds. For instance, is it reasonable to assume that an author's publication rate stays constant through his/her career? Also, isn't it generally accepted that a paper's citation rate varies over time with, at least eventually, a gradual decline? See Burrell (2001, 2002a, 2002b, 2003). And is the basic Poisson model the most appropriate model? What about a negative binomial process for either the productivity or, perhaps more appropriately, the citation process?

Hirsch's (2005) proposal is certainly an interesting idea for a simple index but, as a comparative measure, it would appear from our analysis that its heavy dependence on the underlying parameters means that it cannot be the *single* measure adequate "... to quantify an individual's scientific research output". We concur with Glänzel (2006a) who describes it as "a useful supplement to the bibliometric toolset" but it is certainly not a substitute.

Appendix A

Proof of Theorem 1. Let X_T denote the number of citations achieved by a published paper by time T. For a particular paper published at time $t \in [0,T]$ we have, according to the standard gamma mixture of Poisson processes (GPP) model described by Assumptions 2 and 3,

$$P(X_T = r|t) = \frac{\Gamma(r+\nu)}{r!\Gamma(\nu)}\left(\frac{\alpha}{\alpha+(T-t)}\right)^{\nu}\left(\frac{T-t}{\alpha+(T-t)}\right)^{r} \text{ for } r = 0, 1, 2, \ldots$$

Now, given the total number of publications, say $Y_T = N$, by time T, it is well known that under the assumption that these occur as a Poisson process, the successive publication times $t_1 < t_2 < \ldots < t_N$ are equivalent to the order statistics of a random sample of size N from a uniform distribution on $[0,T]$, see for instance Theorem 2.3.1, p. 67 of Ross (1996). Equivalently, the unordered times are N independent and identically distributed uniform random variables on $[0,T]$, see the Theorem on p. 75 of Stirzaker (2005). Thus, any particular publication time t can be considered to be uniformly distributed on $[0,T]$ so that

$$P(X_T = r) = E_t P(X_T = r|t)\int_0^T \frac{\Gamma(r+\nu)}{r!\Gamma(\nu)}\left(\frac{\alpha}{\alpha+(T-t)}\right)^{\nu}\left(\frac{T-t}{\alpha+(T-t)}\right)^{r}\frac{1}{T}\mathrm{d}t$$

$$= \frac{\alpha^{\nu}\Gamma(r+\nu)}{r!\Gamma(\nu)T}\int_0^T \frac{s^r}{(\alpha+s)^{r+\nu}}\mathrm{d}s \text{ where } s = T-t$$

This can easily be written in terms of a beta distribution of the second kind, see Kleiber and Kotz (2003, Chapter 6). However, if we substitute $y = s/(\alpha+s)$ so that $s = \alpha y/(1-y)$ and $\mathrm{d}s/\mathrm{d}y = \alpha/(1-y)^2$ we find for the integral

$$\int_0^T \frac{s^r}{(\alpha+s)^{r+\nu}}\mathrm{d}s = \int_0^T \left(\frac{s}{\alpha+s}\right)^{r+\nu}\frac{1}{s^{\nu}}\mathrm{d}s = \int_0^{T/(\alpha+T)} y^{r+\nu}\left(\frac{1-y}{\alpha y}\right)\frac{\alpha}{(1-y)^2}\mathrm{d}y = \int_0^{T/(\alpha+T)} \alpha^{1-\nu}y^r(1-y)^{\nu-2}\mathrm{d}y$$

Thus

$$P(X_T = r) = \frac{\alpha\Gamma(r+\nu)}{r!\Gamma(\nu)T}\int_0^{T/(\alpha+T)} y^r(1-y)^{\nu-2}\mathrm{d}y = \frac{\alpha}{(\nu-1)T}\int_0^{T/(\alpha+T)} \frac{\Gamma(r+\nu)}{\Gamma(r+1)\Gamma(\nu-1)}y^{(r+1)-1}(1-y)^{(\nu-1)-1}\mathrm{d}y$$

The integral is now just the cumulative distribution function of a beta distribution with parameters $(r+1)$ and $(\nu-1)$, provided $\nu - 1 \geq 0$ (or $\nu \geq 1$), evaluated at $T/(\alpha+T)$ so that

$$P(X_T = r) = \frac{\alpha}{(\nu-1)T}B\left(\frac{T}{\alpha+T}; r+1, \nu-1\right) \quad \square$$

Proof of Lemma 1. Noting that

$$E[N(n;T)] = E[Y_{\mathrm{T}} P(X_{\mathrm{T}} \geq n)] = E[Y_{\mathrm{T}}]P(X_{\mathrm{T}} \geq n) = \theta T \sum_{r=n}^{\infty} P(X_{\mathrm{T}} = r),$$

the result follows from the above. □

Proof of Lemma 2.

$$\sum_{r=0}^{\infty} B\left(\frac{T}{\alpha+T}; r+1, \nu-1\right) = \sum_{r=0}^{\infty} \int_0^{T/(\alpha+T)} \frac{\Gamma(r+\nu)}{\Gamma(r+1)\Gamma(\nu-1)} y^r (1-y)^{\nu-2} \mathrm{d}y$$

$$= \int_0^{T/(\alpha+T)} \frac{\nu-1}{(1-y)^2} \left(\sum_{r=0}^{\infty} \frac{\Gamma(r+\nu)}{r!\Gamma(\nu)} y^r (1-y)^{\nu}\right) \mathrm{d}y$$

Now the inner summation in this integral expression on the RHS is just the total sum of the probability mass function of a NBD(y, ν) random variable and hence is equal to 1 for any y in [0,1]. Hence, we have that the original sum is equal to

$$(\nu-1) \int_0^{T/(\alpha+T)} \frac{1}{(1-y)^2} \mathrm{d}y = (\nu-1)\left[\frac{1}{(1-y)}\right]_0^{T/(\alpha+T)} = \frac{(\nu-1)T}{\alpha} \quad \square$$

References

Ball, P. (2005). Index aims for fair ranking of scientists. *Nature*, *436*, 900.

Bornmann, L., & Daniel, H.-D. (2005). Does the h-index for ranking of scientists really work? *Scientometrics*, *65*(3), 391–392.

Braun, T., Glänzel, W., & Schubert, A. (2005). A Hirsch-type index for journals. *The Scientist*, *19*(22), 8–10.

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

Burrell, Q. L. (2001). Stochastic modelling of the first citation distribution. *Scientometrics*, *52*, 3–12.

Burrell, Q. L. (2002a). On the *n*th citation distribution and obsolescence. *Scientometrics*, *53*, 309–323.

Burrell, Q. L. (2002b). Will this paper ever be cited? *Journal of the American Society for Information Science and Technology*, *53*, 232–235.

Burrell, Q. L. (2003). Predicting future citation behavior. *Journal of the American Society for Information Science and Technology*, *54*, 372–378.

Cronin, B., & Meho, L. (2006). Using the h-index to rank influential information scientists. *Journal of the American Society for Information Science and Technology*, *57*(9), 1275–1278.

Egghe, L. (2006). An improvement of the H-index: the G-index. *ISSI Newsletter*, *2*(1), 8–9.

Egghe, L. (in press). Dynamic h-index: the Hirsch index in function of time, *Journal of the American Society for Information Science and Technology*.

Egghe, L., & Rousseau, R. (1990). *Introduction to informetrics*. Elsevier: Amsterdam.

Egghe, L., Rousseau, & R. (in press). An informetric model for the h-index. *Scientometrics*.

Glänzel, W. (2006a). On the H-index—a mathematical approach to a new measure of publication activity and citation impact. *Scientometrics*, *67*(2), 315–321.

Glänzel, W. (2006b). On the opportunities and limitations of the H-index. *Science Focus*, *1*(1), 10–11 (in Chinese).

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*(46), 16569–16572. Also available as ar*Xiv:physica/0508113, accessible at http://xxx.arxiv.org/abs/physics/0508025.

Kleiber, C., & Kotz, S. (2003). *Statistical size distributions in economics and actuarial sciences*. New Jersey: Wiley.

Ross, S. (1996). *Stochastic processes* (2nd ed.). New York: John Wiley.

Rousseau, R. (1992). Concentration and diversity of availability and use in information systems. *Journal of the American Society for Information Science*, *43*, 391–395.

Rousseau, R. (2006). A case study: evolution of JASIS' Hirsch index. *Science Focus*, *1*(1), 16–17 (in Chinese).

Stirzaker, D. (2005). *Stochastic processes and models*. Oxford: Oxford University Press.

Van Raan, A. J. (2006). Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgment for 147 chemistry research groups. *Scientometrics*, *67*(3), 491–502.