# Editorial—Quantifying the impact of hydrological studies

## 1.  INTRODUCTION

The two essential elements of the process of scientific publishing are: quality control and dissemination. In recent articles (Kundzewicz & Koutsoyiannis, 2005, 2006), we have examined the quality control of scientific journals achieved by the peer-review system, with particular reference to hydrological journals. The present contribution deals with dissemination and impact of scientific journals, processes which have been greatly affected by the developments of the Internet. Most journals have adapted their procedures and, in addition to the traditional print versions, they provide electronic versions accessible on the Internet. However, this has still not resolved the main dilemma of dissemination of published research results, as the access to journals (except a few open access ones) is restricted to subscribers.

Other ways of publishing research results on the Internet with open and free access, either with or without quality control, have developed in parallel. A recent example of the efficiency of such alternative methods is the case of Grigori Perelman, a Russian mathematician who posted a series of papers with significant results (outline of a proof of the geometrization conjecture, a result that includes the Poincaré conjecture as a particular case) in the open electronic archive arXiv. For these papers, Perelman was awarded the Fields Medal, which is considered to be the highest honour a mathematician can receive (Wikipedia, 2006b). The Perelman papers have been cited already by many and were scrutinized by other mathematicians for their correctness. This case has put into question the practice that impressive scientific publications in prestigious, peer-reviewed journals are a necessary condition of being considered as a candidate for a prestigious scientific award.

Theoretically, open access on the Internet increases accessibility and optimizes dissemination of published scientific results. In practice, however, due to the overload of information on the Internet, it may be difficult to locate and trust a particular article if it is not "authorized" by a reputable body or organization. In this respect, journals "guaranteeing" a minimum quality level of published papers (a level that varies between different journals) have not lost their significant role in the process of dissemination of important scientific results. This is particularly the case in hydrology and water resources, as no reliable alternative to journal publishing has appeared so far. The rising numbers of submitted and published articles in most journals demonstrate that their dissemination role is increasing rather than decreasing. Other evidence of the importance of journals is provided by the fact that, in recent years, several scholarly information systems (see Section 3) have been developed, in which journal articles and citations to them are very important elements.

In *Hydrological Sciences Journal* (*HSJ*) we strive towards improving the dissemination of scientific results. We view our service to the scientific community as a two-stage process. The first stage, the review process, aims at performing quality control, identifying the most important and useful papers, and improving them as much as possible. This is just a preliminary process for the more important second stage, the dissemination of published research results. We see a clear positive feedback between the dissemination and prestige of a journal. Better dissemination increases the popularity and prestige of the journal, so that it attracts more interesting papers, which in

turn enhance dissemination. In recent years, several steps to improve dissemination of *HSJ* have been taken, one of them being the online publication of the journal, in addition to the printed form; other steps are under study.

We have compiled this study to assess the current state in *HSJ*, in comparison to other hydrological journals. In order to do so, we had to explore the entire landscape related to hydrology: articles, authors and journals. We think that the results of the study are useful for the journal and, as (to our knowledge) no other similar study has appeared, we hope that the readers will find it useful too.

## 2.   MEASURES OF DISSEMINATION

An obvious index of dissemination of a journal refers to subscription statistics (number of subscribers—institutional or individual), and assessment of the number of readers (per single subscription). In order to evaluate dissemination of an individual paper in a specific online journal, one can measure the number of visits to (downloads of) a paper (Cowhig, 2003) or to the abstract. The latter is typically openly accessible for any user, whereas access to full papers is typically restricted (to subscribers and pay-per-view clients). This approach is implemented by multi-journal publishers, who arrange compatible counting schemes in all their titles and compare the results. However, the number of downloads (or an index based on this number) is not necessarily a good measure of dissemination and most probably is not related to quality, since the prior expectations raised by the title of the paper and the authors' names may not materialize in the light of the (possibly disappointing) contents of the paper. After downloading a paper, the reader may find it not worthy of interest. Furthermore, such an index would open the door to easy system abuse (e.g. generation of multiple self-downloads). Even the number of Internet links (which is the basis of web pages ranking, e.g. with the PageRank algorithm implemented by Google) to a site which contains the data of a paper cannot provide an objective measure, as indicated by several negative symptoms, such as the so called Google bomb (Wikipedia, 2006a).

A better basic element which may be used to construct objective measures of dissemination is the number of citations to a published article. This is based on solid ground—a paper that fetches many citations deserves respect. So does its author (except in the event of "negative" citations, where many people critically cite a paper to refer to an important error) and so too the journal where it is published and—even more so—a journal containing many papers that are broadly read and quoted. The number of citations measures the popularity of an article, which reflects several factors, such as importance of findings, dissemination mechanism, "crowdedness" of the particular scientific field, and the visibility and prestige of the journal, not to exclude other factors such as "vogue" and persistence (an article that gets cited once is more likely to be accessed and then cited by someone else, etc.). There is also a considerable randomness/chance component. These factors (except the last) seem to be positively correlated to popularity and thus we can use citations to indirectly measure dissemination and importance, having in mind the limitations resulting from the influence of all other factors. Based on the citations, several "bibliometric" indices applied to articles, authors and journals can be constructed, as shown in Box 1.

**Box 1 Bibliometric indices for quantifying research dissemination**

Note: entries with a star are those used in this study.

**For an article**

* Number of citations in other articles (the authors' self-citations may or may not be excluded)

**For an author**

– Total number of articles cited (included in citation data bases)

– Total number of citations for all articles

* h-index: This newly proposed index (Hirsch, 2005) combines the number of articles produced and the number of citations to each for an individual author. An author is assigned an h-index $h$ if there are $h$ of his/her papers that have at least $h$ citations each and all other papers have citations smaller than $h$ each. For instance, if a scientist has published 20 journal papers and the number of citations to particular papers, in decreasing order, read: 9 9 7 6 5 5 3 3 3 2 2 2 1 1 1 1 0 0 0 0, then the value of the h-index is 5.

**For a journal** (see also Fig. 1)

* Impact Factor (IF) of a journal for a specific year ($N$; e.g. 2006) is the quotient $A/B$ where $B$ is the number of articles published in this journal in the two immediately preceding years ($N - 2$, i.e. 2004 and $N - 1$, i.e. 2005) and $A$ is the total number of citations that these articles received during this specific year from other articles in tracked journals (Garfield, 2005). Although the IF can be evaluated based on any database with citations to publications, in practice, cited IF values refer to those determined by ISI.

– Immediacy Index (II) is the average number of times that an article, published in a specific year in a specific journal, is cited over the course of the same year.

– Cited Half-life measures the number of years, going back in time from the current year, that account for half of the total number of citations received by the cited journal in the current year.

* Citation density (CD) of a journal for a specific year is the mean number of citations received per article published in the journal in the specific year, i.e. the quotient $R/S$ where $S$ is the total number of articles published in this journal in this specific year of concern and $R$ is the cumulative number of citations that these articles received over the following years until present (or over a specified period). Note that some papers continue to be cited for a long time after publication.

Most indices in Box 1 are widely known. The graphical interpretation of the notion of Immediacy Index, the Impact Factor, and the Cited Half-life for a schematic temporal distribution of citations is presented in Fig. 1. Interestingly, the two curves of Fig. 1 have properties familiar to hydrologists, as they are reminiscent of the response in linear rainfall–runoff models. Very recent articles are not cited very often because the dissemination period has not lasted long, so that they are not well known. Thus, the number of citations per year increases with the number of years from publication, but, after a certain number of years, papers become outdated and the citations to them become less frequent, so at this stage the number of citations decreases with the time distance from present.
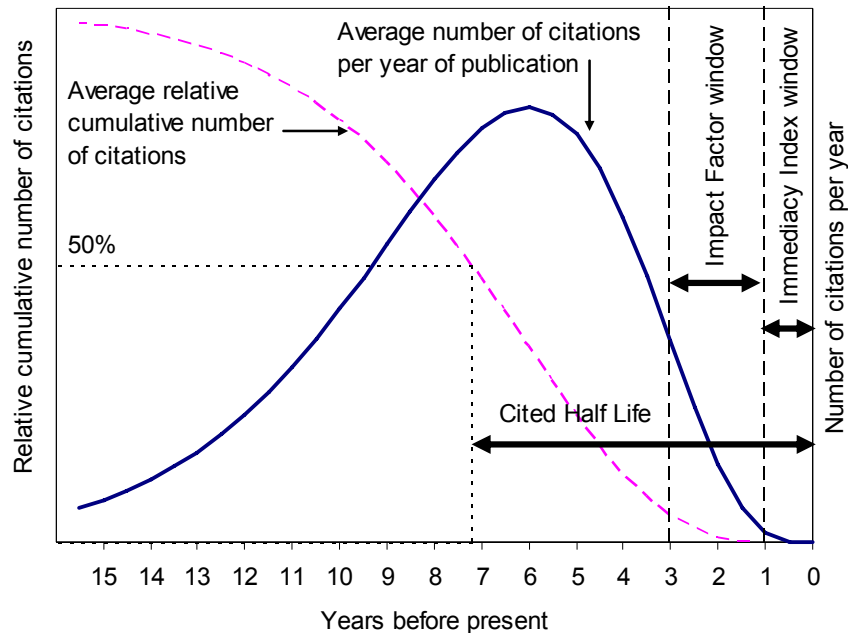
**Fig. 1** Explanation sketch for the definitions of bibliometric indices for a journal; time zero should read as 31 December of the present year. In the presented example, given for illustration only, papers published 6–7 years ago receive maximum number of citations in the present year.

The CD index for a journal is not as well known as IF, but this index of dissemination is potentially more useful for a researcher than IF or other indices. Certainly, a researcher is interested in whether his/her paper becomes known and receives citations soon after its publication (contributing to the II and IF indices). But even more interesting is the total number of citations in the entire time span from the moment of publication until the present, and this is better encapsulated by the CD.

The h-index for an author is less well-known, as it was proposed very recently. This interesting, simple and easily calculated index was adopted in this study; we expect that this index will be used increasingly. It is an excellent metric for the evaluation of an individual scientist as it focuses on the most important achievements disregarding less important ones (the latter being perhaps driven by the "publish or perish" syndrome). An attractive property of the h-index is its cumulative nature. Its value does not decrease with time; it can only grow, or at least remain stable: what has been achieved remains recognized forever. Note for comparison that the IF used for evaluation of journals refers to a limited time window of two years. Hence, if a journal did very well in the past but does not attract many new citations, its IF falls in a merciless way.

There are obvious positive correlations between different bibliometric indices of the same category: for instance the h-index is positively (but not necessarily strongly) correlated to the total number of citations to publications of an author and CD is positively correlated to IF.

There are also indirect links between different indices referring to articles, authors and journals which are responsible for several positive feedback mechanisms. For instance, an article by a well-known author who has a high h-index will most probably receive more citations than a similar article of comparable quality, published by an

unknown author. The rule "*noblesse oblige*" drives scientists to expect high quality, and hence to read (and cite) works of an established author. A researcher tends to choose a journal with a high IF to publish his/her best papers and thus enhances citations to the journal. Also, a researcher or a student will (likely) prefer journals with higher IF in his/her bibliographic search. A librarian will select journals with highest IF for the collection, thus improving the dissemination of these journals. These positive feedbacks are known to researchers and this amplifies the feedbacks. For instance, an article by a particular author in a very reputable journal is typically regarded as more important than an article by the same author published in a less reputable journal. For this reason, the reputability of the journals where a researcher has published articles affects his/her own reputation. However, sometimes there are other factors affecting the choice, such as the opportunity created by a special issue / special section, manuscript processing and publication speed, or inter-personal relations. Some scientists regard the process of selection of a journal for submission of a paper as an optimization exercise. If a paper is smoothly accepted in a journal $X$ with impact factor $IF_X$, the author may not be satisfied, because the paper might have been accepted in a journal $Y$, with the impact factor $IF_Y > IF_X$.

Typically, citations to a research article originate from other research articles published in scientific journals. This should be regarded as a weakness because, in several applied fields, including hydrology, the importance of a research result is obviously related to its adoption and assimilation into professional practice, resulting in reports rather than research articles. However, more recently developed databases include also citations to other sources, e.g. reports, provided that these sources are accessible on the Internet.

## 3. DATA AND METHODOLOGY

The most widely known source of bibliometric information is the Institute for Scientific Information—ISI—in Philadelphia, USA. Recently, other database systems with similar information have been developed, among which Scopus (an initiative of the Elsevier publishing company) and Google Scholar (by Google) are general purpose sources; there exist also discipline-specific sources, but these have little or no relation to hydrology. Of these three general purpose sources, Google Scholar is open access (still operating in a Beta version), whereas the other require subscriptions for access. Here, we used all three sources to retrieve and compile raw data for our analyses. In addition, we used already processed results (particularly from Journal Citation Reports issued by ISI). Table 1 shows what we see as the advantages and disadvantages of the three data sources.

Comprehensive information about all the periodicals tracked by ISI (over 43 000 journal titles), contained in Garfield (2005), was used as background information for this study. However, as our focus was on water resources and hydrological papers, we retrieved and processed data from the above mentioned sources for a selected set of such journals, as shown in Table 2. The ISI 2005 Journal Citation Report, includes 57 journals in the Water Resources category, some of which are commonly regarded as general hydrological, whereas others are focused on different water resources disci-plines (e.g. water quality) and even though they are related to hydrology, their particular specialization would classify them in different sub-categories. We found it

**Table 1** Data sources used in this study and our evaluation of their main advantages and disadvantages.

| Data source | Advantages | Disadvantages |
|---|---|---|
| ISI Web of Science (http://www.isinet.com) | Complete data base<br>Prompt updates | Difficult/inflexible handling<br>High cost |
| Scopus (http://www.scopus.com) | Easy and flexible search<br>Handling of excessive records<br>Link to source | Missing data (for old papers)<br>No citations prior to 1996 |
| Google Scholar (http://scholar.google.com) | Free access<br>Large coverage<br>Link to source | Limited user control<br>Infrequent updating |

**Table 2** Groups of water resources journals used in the study (out of a total of 57 water resources journals included in ISI 2005 Journal Citation Report, of which 12 are shown).

| Journal abbreviation | ISI IF, 2005 | ISI rank in water resources category, 2005 | Top 5 water resources journals | Top 3 hydrological journals | Top 5 hydrological journals |
|---|---|---|---|---|---|
| *Water Res.* | 3.019 | 1 | ✓ | | |
| *Water Resour. Res.* | 1.939 | 2 | ✓ | ✓ | ✓ |
| *Aquat. Conserv.* | 1.833 | 3 | | | |
| *J. Hydrol.* | 1.745 | 4 | ✓ | ✓ | ✓ |
| *J. Contam. Hydrol.* | 1.733 | 5 | ✓ | | |
| *Adv. Water Resour.* | 1.684 | 6 | ✓ | ✓ | ✓ |
| *Hydrol. Sci. J.* | 1.606 | 7 | | | ✓ |
| *Irrigation Sci.* | 1.605 | 8 | | | |
| *River Res. Appl.* | 1.426 | 9 | | | |
| *Ground Water* | 1.419 | 10 | | | |
| *Clay Clay Miner.* | 1.364 | 11 | | | |
| *Hydrol. Process.* | 1.336 | 12 | | | ✓ |

very difficult, in the framework of this study, to collect and process data for all 57 Water Resources category journals. Rather, we chose seven journals using three criteria: (a) high IF, (b) relevance to hydrology, and (c) long period of journal's presence in the ISI Water Resources category.

We re-arranged the seven journals forming three groups, as shown in Table 2. For convenient orientation we labelled the groups as "top 3" or "top 5" but this is somewhat arbitrary and misleading because, in fact, journal ranking based on IF varies in different years; however, we were cautious not to include in these groups journals with high variation of ranking. As a rule of thumb (Amin & Mabe, 2000), journals with impact factors that differ by less than 25% belong to the same rank. The retrieved data are for the publication period 1982–2005, which is the time span of *HSJ* under its present title. *Hydrological Sciences Journal* actually spans a much longer period of 51 years, being the oldest among hydrological journals, but its title has changed twice: initially—*Bulletin of the International Association of Scientific Hydrology* (1956–1971); and later—*Hydrological Sciences Bulletin* (1972–1981). Six of the journals cover the whole study period, whereas *Hydrological Processes* was launched in 1987. The period of retrieved citations was 1996–2005 when the data origin was Scopus, but occasionally larger periods were considered for other sources. The methodological (probabilistic and statistical) background used for processing and interpreting the data is given in the Appendix.

## 4. ARTICLES

The most prominent characteristic of the number of citations received by individual articles is the enormous variability. According to Garfield (2005), among the 38 163 319 articles tracked by ISI (as of August 2005), 47.8% have not received any citation at all, while the record highest number of citations to a single article is 293 328 (for a biological article by Lowry *et al*., 1951). The probability distribution of citations for the population of all articles is shown in Fig. 2. With the explanations given in the Appendix, it may be concluded that the variance of the distribution diverges (it would be infinite for a population size tending to infinity; note that the slope of the survival function in the double logarithmic plot is about –2).

It is well known that different scholarly disciplines have very different populations, and publication and citation practises, which affect the number and the accumulation rate of citations. Therefore, it is more relevant to consider the citations to an article in a category of its peers, rather than within all tracked journals. Disciplines with high rates of citations are the biological and medical sciences and physics. Hydrological articles generally have much lower citation rates. The probability distribution for articles published in the top five hydrological journals is shown in Fig. 2. The apparent differences from that of the case of all journals are the following:
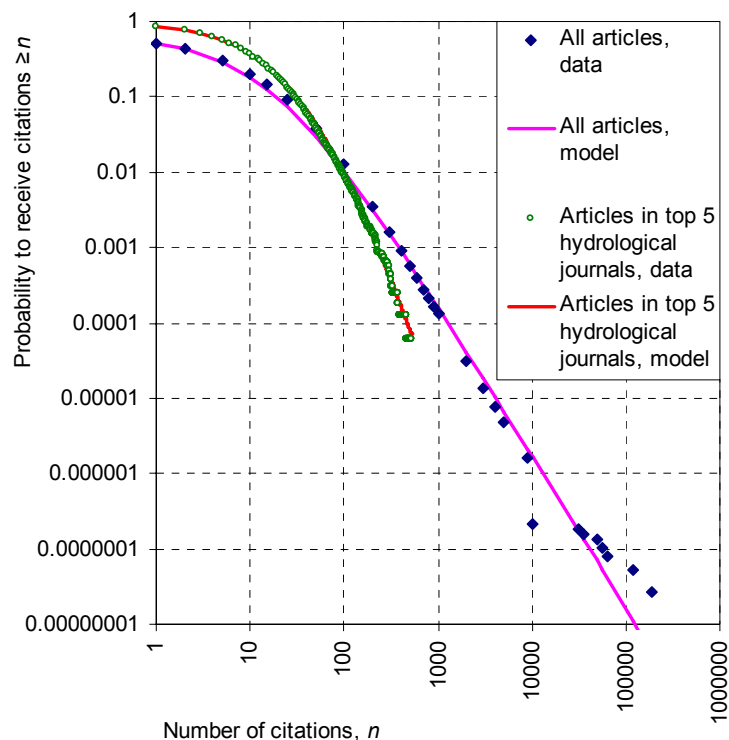


**Fig. 2** Probability distribution of the number of citations to a journal article in general and an article published in one of the top five hydrological journals; "model" is equation (A1) (Appendix) in its four-parameter form for a general article with $\kappa = 0.5$, $\theta = 0.25$; $\lambda_0 = 0.562$, $\lambda_1 = 0.259$ and $\lambda_2 = 1.651$, and in its three-parameter form for a hydrological article with $\kappa = 0.2$, $\theta = 0.81$; $\lambda_0 = 2.05$, $\lambda_1 = 0$ and $\lambda_2 = 0.316$. Data origin: ISI from Garfield (2005) for general articles and Scopus for hydrological articles.

(a) the variability is lower (as expected because of considering a single discipline and the relatively low difference of quality between the leading journals in the field);

(b) the frequency of zero citations is significantly lower (12.8%; because of the fact that the journals considered are the most prestigious in the area; as opposed to about 50% in the large population of all journals, which includes the "outsiders" that usually fetch no citations); and

(c) the frequency of high numbers of citations is also significantly lower (because the hydrological sciences are not densely populated).

According to the results of our search (which was based on repeated mining of data from the three databases using several keywords such as "water", "hydrology" etc.), the record number of citations in hydrological sciences slightly exceeds a thousand. In our search we located several articles with hydrological content that have more that 500 citations, as listed in Table 3. It is interesting to note that they have been published in top hydrological journals (with *Water Resour. Res.* prevailing and including *Hydrol. Sci. Bull.*). Some of these seminal papers, as well as other, less cited benchmark papers are now being re-visited in a recent IAHS initiative to re-publish important historical papers in hydrology in a series of thematic volumes, the first of which has already appeared (Beven, 2006). We also searched highly cited articles with content related to hydrological sciences that were published in non-hydrological journals. These are also shown in Table 3. Interestingly, they come from journals in related disciplines (except for the old paper by Penman, 1948, which pre-dates the establishment of disciplinary journals), rather than from highly cited, broad-scope journals such as *Science* and *Nature* (see Section 6). As shown in Table 3, all three sources of information were used to identify these articles. In many cases, there is

**Table 3** Some of the most popular hydrological papers that have received more than 500 citations.

| Authors and year of publication | Paper subject | Journal abbreviation | No. of citations* |
|---|---|---|---|
| *Hydrological papers in Water Resources journals:* | | | |
| Nash & Sutcliffe (1970) | Conceptual models | *J. Hydrol.* | 1140 / 1122 / 886 |
| Topp *et al*. (1980) | Soil water content | *Water Resour. Res.* | 1052 / 945 / 880 |
| Mualem (1976) | Unsaturated flow | *Water Resour. Res.* | 1003 / 773 / 747 |
| Beven & Kirkby (1979) | Contributing area model | *Hydrol. Sci. B.* | 844 / 840 / 799 |
| Beven & Germann (1982) | Flow in soils | *Water Resour. Res.* | 785 / 519 / 447 |
| Gelhar & Axness (1983) | Macrodispersion in aquifers | *Water Resour. Res.* | 784 / 465 / 313 |
| Ritchie (1972) | Crop evaporation | *Water Resour. Res.* | 708 / 393 / 304 |
| Clapp & Hornberger (1978) | Soil hydraulics | *Water Resour. Res.* | 704 / 500 / 389 |
| *Papers with relevance to hydrology in non-Water Resources journals:* | | | |
| van Genuchten (1980) | Hydraulic conductivity | *Soil Sci. Soc. Am. J.* | 2604 / 2276 / 1054 |
| Penman (1948) | Evaporation | *Proc. Royal Soc. Lond.* | 1556 / 778 / 769 |
| Mandelbrot & van Ness (1968) | Fractional Brownian noise | *SIAM Review* | 1258 / 966 / 1366 |
| Milliman & Meade (1983) | Sediment delivery | *J. Geol.* | 779 / 613 / 449 |
| Hurst (1951) | Long-term persistence | *Trans. ASCE* | 726 / 636 / 803 |
| Xie & Arkin (1997) | Global precipitation | *B. Am. Meteorol. Soc.* | 654 / 687 / 647 |
| Hosking (1981) | Fractional differencing | *Biometrika* | 523 / 487 / 868 |

* ISI / Scopus / Google Scholar.

significant disagreement between the number of citations in the different databases, perhaps due to missing data in databases, human errors in data entry, spelling problems (especially if authors have long and difficult names, e.g. those of the authors of this article), and algorithmic defects. For these and other reasons our list in Table 3 may be incomplete, but we look forward to feedback (formal or informal commentaries) to make it more complete.

## 5. RESEARCHERS/AUTHORS

The bibliometric study of outstanding scientists' works makes it possible to illustrate how relevant the scientific community regards such types of indices to assess the overall contribution of an individual. To make a compass record of outstanding hydrologists, we used the list of scientists awarded with the International Hydrological Prize (IHP) by the International Association of Hydrological Sciences (IAHS). The list of laureates shown in http://www.cig.ensmp.fr/~iahs/IHPrize/IHPWinners.htm includes 27 hydrologists (1981–2006). The IHP is awarded to persons who have made an outstanding contribution to hydrology, with an international dimension extending beyond both the country of normal work and the specific field of interest of each scientist. The contribution may have been made through scientific work, as demonstrated by publication in international journals of scientific literature of a high standard, or through practical work, as demonstrated by reports of the projects concerned, or (preferably) through both scientific and practical work. This two-component definition of the IHP makes the list of its laureates ideal for the purpose of our study, which is not to do any type of comparison but to assess the relevance of bibliometric indices with the contribution of scientists (assuming that the significant contribution of IHP laureates is not disputable).

The h-index was determined for each IHP laureate; the results of the investigation are depicted in Fig. 3. Again, the record (Fig. 3(a)) exhibits a strong variability of the h-index (where even the mean is diverging according to the framework in the Appendix). There is a large (25%) frequency of zero h-index (seven laureates, five awarded in the first decade of the Prize, in the 1980s) which signifies the fact that people with no cited publications can contribute greatly to hydrology (e.g. in hydrological practice). From Fig. 3(b) we can distinguish two periods, before and after 1990. It seems that, in the first period, the number of citations to articles was not regarded as an important criterion of recognition of the contribution, but this may have changed in the second period. The slightly decreasing rate after 1995 is perhaps an artefact of the accumulation character of the h-index (which typically grows with the age of a researcher). Until 1993, 11 out of the 15 laureates have an h-index below 2. This perhaps illustrates that the "tyranny of impact factors" and other bibliometric indices (Colquhoun, 2003) is relatively recent. The data shown in Fig. 3 may seem quite surprising to the younger readership. In the past, publishing journal articles in "top notch" journals was not considered necessary. It was more important in many countries to write (or edit) books, and to publish in national journals (often not included in databases). Another observation from Fig. 3 (c) and (d) is that scientists from Europe have a smaller h-index than those from other continents, which may reflect the fact that in Europe the contribution of a scientist has not been so much determined in terms of
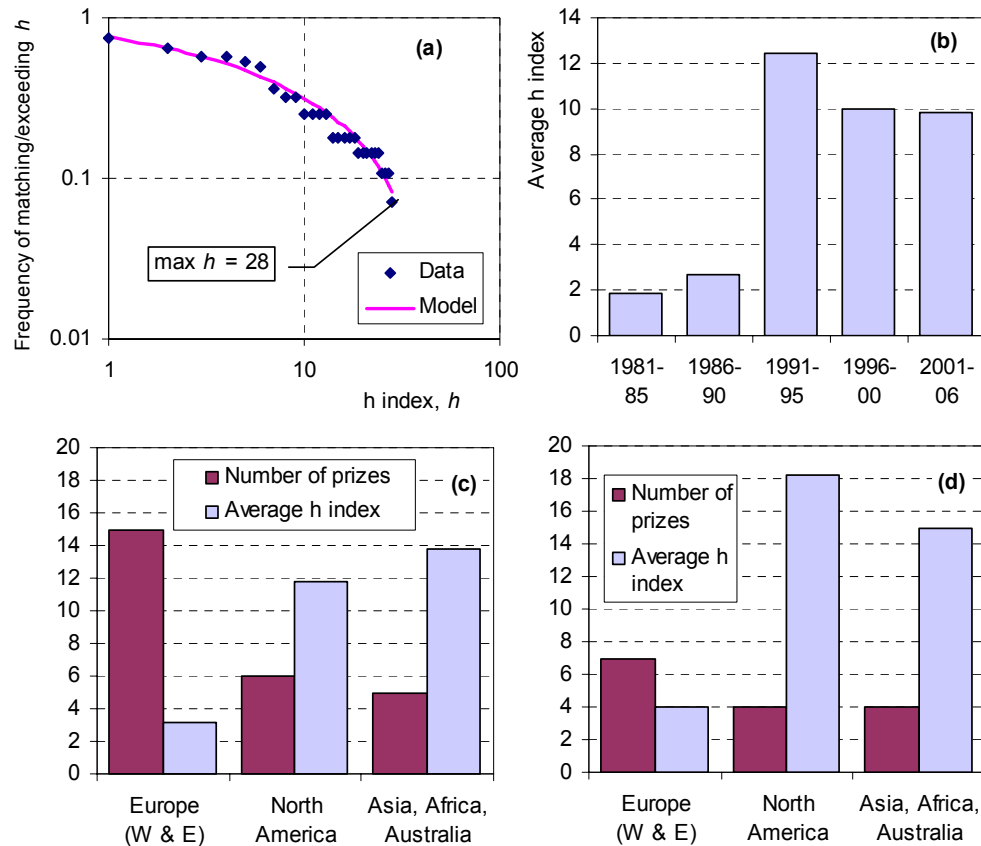
**Fig. 3** Statistical depiction of h-index for scientists that have received IAHS's International Hydrology Prize: (a) frequency distribution; (b) temporal distribution; (c) geographical distribution through the whole 25-year period (1981–2006); (d) geographical distribution for the 15-year period (1991–2006). "Model" in (a) is equation (A1) in its two-parameter form with $\kappa = 50$, $\lambda_0 = 0.0645$ and $\lambda_1 = 0.0785$. Data from ISI; note that the time base extends up to now and not up to the year of receiving the Prize.

publications and indices. It should be also recalled that, in the case of scientists from Eastern Europe, up to the late 1980s, attending international conferences abroad and delivering presentations there was deemed more important (and more difficult, under the political and economic system) than publishing articles.

The highest h-index among the 27 awarded hydrologists is 28 (the late J. R. Philip, awarded the IHP in 1995), while the value of 20 was also exceeded by W. J. Shuttleworth (27), 2006 IHP winner and P. S. Eagleson (24), 1991 IHP winner. Our search located hydrologists with h-index higher than 30 (obviously, we do not provide a complete list, as we may have missed some names): K. Beven (41), I. Rodriguez-Iturbe (39), M. T. van Genuchten (37), S. P. Neumann (36), L. W. Gelhar (35), G. M. Hornberger (31) and J. T. Ritchie (31). It is striking that this list is domi-nated by groundwater scientists. Note, for comparison, that the physicist and biologist with the highest h-index are E. Witten (110) and S. H. Snyder (191), respectively (Hirsch, 2005; Wikipedia, 2006c).

## 6.  JOURNALS

The ISI IF, being the most frequently used index for ranking journal quality, has numerous shortcomings. It strongly varies with the subject field (cf. fundamental *vs* applied sciences), size of the journal (small-scale journals have large IF fluctuations), number of authors (statistically, multiple authorship tends to boost the IF value); cf. Amin & Mabe (2000). The index is only available for those journals that ISI decides to cover—most national scientific journals are not included and a language barrier exists (e.g. for Chinese or Japanese); the many other types of publication and research reporting are not covered. The time window of two years is unlikely to be optimal for several, slow-reaction areas (where a five-year window could be better). However, the index is in common use in journal evaluation, helping librarians in decision making, publishers in monitoring journals, editors in assessing effectiveness of policies, as well as in research assessment exercises.

As already discussed and also shown in Table 4, the journals with the highest IF are from the medical and biological sciences and physics. Table 4 also includes two broad-scope journals, *Science* and *Nature*, which are the most reputable among all scientific disciplines.

Journals in the water resources and hydrology fields have IF lower than those of Table 4 by more than an order of magnitude (Table 2). The evolution of the IF of two groups of journals (top five water resources journals and top three hydrology journals) in the last 13 years is shown in Fig. 4. The gradual increase may manifest the intensi-fication of research and scientific publications in recent years and the increasing interest of the scientific community in objective bibliometric indices, as discussed already.

A diagram for the CD of the top five hydrological journals and for a longer period (24 years) is shown in Fig. 5. To estimate the CD values, we used raw data from Scopus. To interpret this diagram, one must bear in mind the cumulative character of CD. For instance, the value CD = 18.8 in 1995 is the average number of citations received in the period 1995–2005 per article published in 1995, whereas the value 11.2 in 2000 is the average number of citations received in the period 2000–2005 per article published in 2000. Thus, in a "stationary" system CD would be a decreasing function of time: older papers tend to have (on average) more citations than newer ones, because more citations are added as time goes by. The rising limb prior to 1993, which contrasts this property, is explained by the fewer publications and citations in this period, in comparison to the last decade, and the missing citations in Scopus before 1996.

**Table 4** Journals with IF above 25 according to ISI 2005 Journal Citation Report.

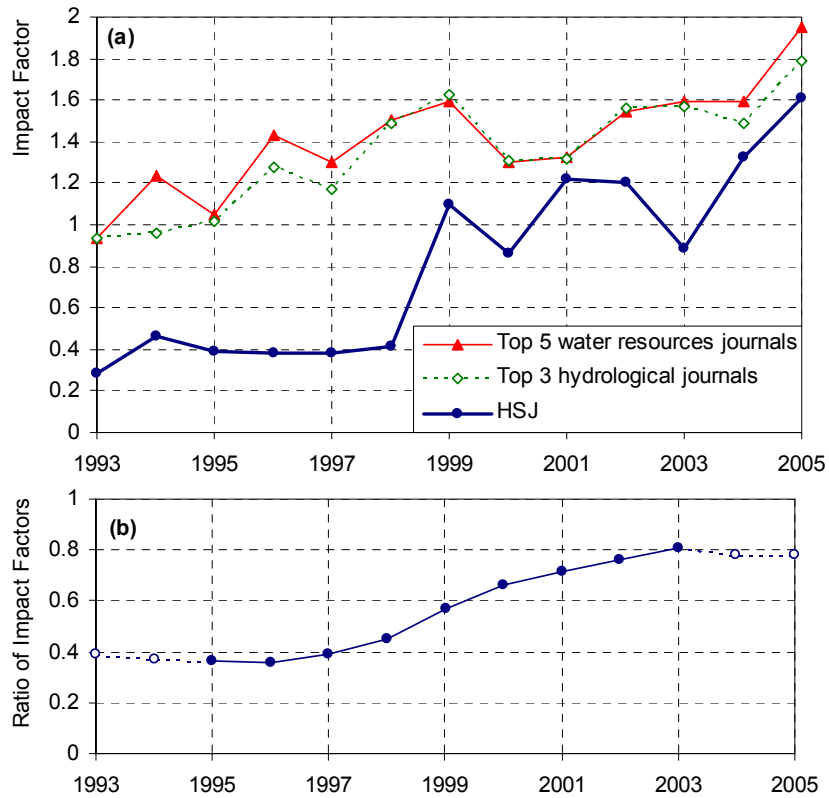| Rank | Journal abbreviation | IF | Rank | Journal abbreviation | IF |
|------|----------------------|-----|------|----------------------|-----|
| 1 | *Ca-Cancer J. Clin.* | 49.794 | 9 | *Nat. Rev. Mol. Cell Bio.* | 29.852 |
| 2 | *Ann. Rev. Immunol.* | 47.400 | 10 | *Cell* | 29.431 |
| 3 | *New Engl. J. Med.* | 44.016 | 11 | *Nature* | 29.273 |
| 4 | *Ann. Rev. Biochem.* | 33.456 | 12 | *Nat. Med.* | 28.878 |
| 5 | *Nat. Rev. Cancer* | 31.694 | 13 | *Physiol. Rev.* | 28.721 |
| 6 | *Science* | 30.927 | 14 | *Nat. Immunol.* | 27.011 |
| 7 | *Nat. Rev. Immunol.* | 30.458 | 15 | *Nat. Genet.* | 25.797 |
| 8 | *Rev. Mod. Phys.* | 30.254 | | | |

**Fig. 4** (a) Evolution of the impact factors averaged over the top five water resources journals and over the top three hydrological journals, in comparison with the impact factor of *HSJ*; (b) evolution of the IF of *HSJ* with respect to that of top three hydrological journals, smoothed by taking (centred) 5-year moving averages (except two points in each edge that are averages of 4 or 3 years). Data from ISI Journal Citation Reports.
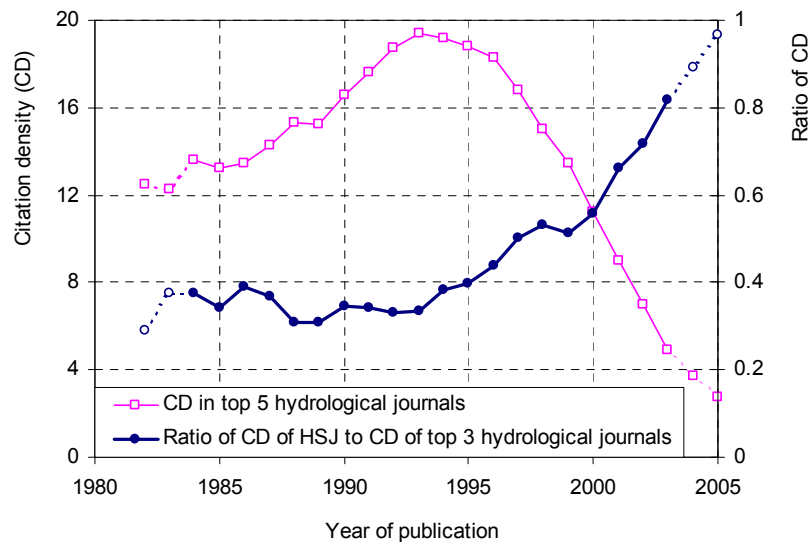


**Fig. 5** Average CD for articles published in the top five hydrological journals *vs* the year of publication and evolution of the ratio of CD of the *HSJ* papers to that of top three hydrological journals. The citation period examined is from 1996 to August 2006. To smooth out random fluctuations, both curves are centred 5-year moving averages except two points in each edge that are averages of 4 or 3 years. Data from Scopus.

In both Figs 4 and 5 we have plotted, separately, the index of *HSJ* (IF and CD) as a ratio of the value for the top three; both indicate a spectacular improvement. The IF of *HSJ* jumped from about 0.4 in 1998 and before, to 1.606 (four times higher) in 2005. The CD was almost constantly below 40% of the average of the top three hydrological journals before 1995, but today this percentage exceeds 95%. We hope that this will be appreciated by contributors to *HSJ*. On average, papers published in *HSJ* are now anticipated to receive nearly as many citations as if they were published in one of top three hydrological journals. An unfortunate consequence of the improvement in *HSJ* indices is that, due to the significant increase of the number of papers submitted to *HSJ,* the rejection rate for incoming papers has increased.

## 7.  CONCLUSIONS AND DISCUSSION

– The number of citations to published articles provides a satisfactory (but not ideal) basis to characterize their dissemination and impact.
– Citations to hydrological papers are much less frequent than in more populated disciplines (e.g. medicine); yet some hydrological papers have gathered very many citations.
– The highest ranked hydrological papers in terms of citations come mostly from top hydrological journals or from journals in related disciplines, rather than from broad-interest interdisciplinary journals.
– In the hydrological community, bibliometric indices have gained more interest and respect in the last decade. However, the community still does not regard any of these indices as a decisive criterion. Other criteria to assess the contribution of a scientist are deemed important, albeit difficult to quantify.
– Generally, the quantification of the quality and importance (which are not identical to popularity) of research achievements is very difficult, if not impossible. Some papers which fetch zero citations may be very important and of high quality. This may happen to papers reporting research ahead of their time, being beyond the state of the art and thus not recognized currently.
– *Hydrological Sciences Journal* has exhibited spectacular progress in terms of bibliometric indices, which now classifies it in the group of top hydrological journals and demonstrates that it offers high dissemination opportunity to published articles.
– However, the IF is not the only factor that matters in *HSJ*, as we work in a multi-objective setting. Quality concerns come first in *HSJ*, but this is an IAHS journal, that has to support hydrology and hydrologists in less developed countries. Hence, the role of mentor is often assumed by the journal's editorial board and referees. Moreover, the statutory bi-lingual character of the Journal is not advantageous from the IF viewpoint: our search revealed that most of *HSJ* papers written in French do not fetch any citation and only a few of them fetch a single citation. Thus, our francophone authors who are interested in the dissemination of their results should perhaps think twice before deciding to write their paper in French.

## REFERENCES

Amin, M. & Mabe, M. (2000) Impact factors: use and abuse. *Perspectives in Publishing* (Elsevier) **1**, 1–6.

Beven, K. J. (ed.) (2006) *Streamflow Generation Processes*. IAHS Benchmark Papers in Hydrology Series (Series Editor J. J. McDonnell), IAHS Press, Wallingford, UK.

Beven, K. & Germann, P. (1982) Macropores and water flow in soils. *Water Resour. Res.* **18**(5), 1311–1325.

Beven, K. J. & Kirkby, M. J. (1979) A physically based, variable contributing area model of basin hydrology. *Hydrol. Sci. Bull.* **24**(1), 43–69.

Clapp, R. B. & Hornberger, G. M. (1978) Empirical equations for some soil hydraulic properties. *Water Resour. Res.* **14**(4), 601–604.

Colquhoun, D. (2003) Challenging the tyranny of impact factors. *Nature* **423** (29 May 2003), 479.

Cowhig, J. (2003) Measuring the "impact" of a scientific paper. *Chinese Journal of Scientific and Technical Periodical* **14**(2), 117–120.

Garfield, E. (2005) The agony and the ecstasy—The history and meaning of the journal Impact Factor. International Congress on Peer Review and Biomedical Publication, Chicago (16 September 2005), (http://www.garfield.library.upenn.edu/papers/jifchicago2005.pdf).

Gelhar, L. W. & Axness, C. L. (1983) 3-dimensional stochastic-analysis of macrodispersion in aquifers. *Water Resour. Res.* **19**(1), 161–180 .

Hirsch, J. E. (2005) An index to quantify an individual's scientific research output. *Proc. Nat. Acad. Sci. USA* **102**(46), 16569–16572.

Hosking, J. R. M. (1981), Fractional differencing. *Biometrika* **68**(1), 165–176.

Hurst, H. E. (1951) Long term storage capacities of reservoirs. *Trans. Am. Soc. Civil Engrs* **116**, 776–808.

Koutsoyiannis, D. (2005) Uncertainty, entropy, scaling and hydrological stochastics, 1, Marginal distributional properties of hydrological processes and state scaling. *Hydrol. Sci. J.* **50**(3), 381–404.

Kundzewicz, Z. W. & Koutsoyiannis, D. (2005) Editorial—The peer-review system: prospects and challenges. *Hydrol. Sci. J.* **50**(4), 577–590.

Kundzewicz, Z. W. & Koutsoyiannis, D. (2006) Pathologies, improvements and optimism. *Hydrol. Sci. J.* **51**(2), 357–363.

Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall R. J. (1951) Protein measurement with the Folin phenol regent. *J. Biol. Chem.* **193**, 265–275.

Mandelbrot, B. B. & van Ness, J. W. (1968) Fractional Brownian motions, fractional noises and applications. *SIAM Review* **10**(4), 422–437.

Milliman, J. D. & Meade, R. H. (1983) World-wide delivery of sediment to the oceans. *J. Geol.* **91**(1), 1–21.

Mualem, Y. (1976) New model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resour. Res.* **12**(3), 513–522.

Nash, J. E. & Sutcliffe, J. V. (1970) River flow forecasting through conceptual models, part I, A discussion of principles. *J. Hydrol.* **10**(3), 282–290.

Penman, H. L. (1948) Natural evaporation from open water, bare soil and grass, *Proc. Roy. Soc. Lond. Ser. A, Math. and Phys. Sci.* **193**(1032), 120–145.

Ritchie, J. T. (1972) Model for predicting evaporation from a row crop with incomplete cover. *Water Resour. Res.* **8**(5), 1204–1213.

Topp, G. C., Davis, J. L. & Annan, A. P. (1980) Electromagnetic determination of soil water content: measurements in coaxial transmission lines. *Water Resour. Res.* **16**(3), 574–582.

Tsallis, C. (1988) Possible generalization of Boltzmann-Gibbs statistics. *J. Statist. Phys.* **52**, 479–487.

van Genuchten, M. Th. (1980) Closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.* **44**(5), 892–898.

Xie, P. P. & Arkin, P. A. (1997) Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Am. Met. Soc.* **78**(11), 2539–2558.

Wikipedia contributors (2006a) Google bomb, *Wikipedia, The Free Encyclopedia;* http://en.wikipedia.org/wiki/Google_bomb (as of 3 September 2006).

Wikipedia contributors (2006b) Grigori Perelman, *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Grigori_Perelman (as of 3 September 2006).

Wikipedia contributors (2006c) H-index, *Wikipedia, The Free Encyclopedia;* http://en.wikipedia.org/wiki/H-index (as of 3 September 2006).

## APPENDIX

### Statistical properties and modelling of bibliometric indices

Bibliometric indices are generally variables of discrete type. To perform statistical analyses and interpretation of statistical behaviours we may think each of them as a discrete random variable $M$ having a probability mass function $p_m := P\{M = m\}$ where

*P* denotes probability. A given population of a bibliometric index (e.g. the number of citations of each of the articles included in a database) is usually large but finite. The use of the term probability for this specific population, instead of the more precise—for this case—term frequency, signifies an abstraction, a cognitive replacement of the actual population with a hypothetical mathematical model. In this model it may be sometimes convenient to think of *M* as an unbounded variable, even though in a real population it is bounded.

A convenient and flexible model for the probability mass function of *M* is:

$$p_m = [1 + \kappa\,(\lambda_0 + \lambda_1\,m + \lambda_2\,m^{\theta})]^{-1-1/\kappa} \tag{A1}$$

This is a generalized distribution with four parameters, namely $\kappa$, $\theta$, $\lambda_1$ and $\lambda_2$, whereas $\lambda_0$ is not a free parameter but an adjustable coefficient whose value is determined so that the sum of all probabilities equals 1. The rationale and derivation of this distribution (with $\theta = 2$) is based on the maximization of Tsallis (1988) entropy and can be found in Koutsoyiannis (2005)—albeit for modelling of hydrological variables. Interesting special forms of the distribution are the following:

- A single parameter form is obtained when $\kappa = 0$ and $\lambda_2 = 0$ ($\theta$ is indifferent); application of de l'Hôpital's rule results in this case in $p_m = \exp(\lambda_0 + \lambda_1\,m)$, which is the geometrical distribution.
- A two parameter form is obtained when $\kappa \neq 0$ and $\lambda_2 = 0$; it is easily seen that this is the well-known Zipf-Mandelbrot law.
- A three parameter form is obtained when $\lambda_1 = 0$.

For $\kappa > 0$ this distribution has a long (power-law) tail, which indicates high variability. For an unbounded *M*, the *r*th moment of the distribution diverges if $r \geq r^*$, where $r^* = 1/\kappa$ when $\lambda_2 = 0$ or $\theta \leq 1$ and $r^* = \theta/\kappa + \theta - 1$ otherwise. In the first case, which is most frequent, a value $\kappa \geq 0.5$ indicates that the variance does not exist (it is infinite) whereas a value $\kappa \geq 1$ indicates that even the mean does not exist. As in reality populations are finite, we can always calculate the mean and variance but we must bear in mind that these statistics may not converge as the population size increases and thus, they are quite sensitive quantities, not well defined.

**Demetris Koutsoyiannis** *(Deputy Editor)*
**Zbigniew W. Kundzewicz** *(Editor)*