BULLETIN OF THE POLISH ACADEMY OF SCIENCES
TECHNICAL SCIENCES, Vol. 61, No. 1, 2013
*DOI: 10.2478/bpasts-2013-0001*

INVITED PAPERS

# Doing Hirsch proud; shaping H-index in engineering sciences

L. CZARNECKI[1]*, M.P. KAŹMIERKOWSKI[2], and A. ROGALSKI[3]

[1] Building Research Institute, 1 Filtrowa St., 00-611 Warsaw, Poland
Faculty of Civil Engineering, Warsaw University of Technology, 16 Armii Ludowej Ave., 00-637 Warsaw, Poland
[2] Institute of Control and Industrial Electronics, Warsaw University of Technology, 75 Koszykowa St., 00-662 Warsaw
[3] Institute of Applied Physics, Military University of Technology, 2 Kaliskiego St., 00-908 Warsaw, Poland

**Abstract.** The h-index concept has been analysed in aspects of a contemporary tendency of parameterisation of everything and as the potential measure of the knowledge progress, which recognises individuals, institutions and Engineering sub-disciplines that best generate new knowledge. Considerations have been presented at the level of universality of knowledge which implies permanent progress and on the base of careful thoughts of the domestic experience. The h-nature of things has been described, and several axiomatic characterisations of the Hirsch index have been gathered. The mechanism how to increase the h-index has been presented. Some similarities between h-index and the journal impact factor (JIF) have been stressed. Also the universal role of H-index in ranking countries in all areas and in Engineering has been exampled in extended tables.

The Glänzel's model which connects the h-index with two fundamental scientometric indicators: number of publications and the rate of citation, has been analysed. Following the Microsoft Academic Search, the lists of 15 top scientists from various academic disciplines and separately in Engineering have been composed. It has been found that the population of the best keeps basically the same relations between the h-index and a number of publication, and between the h-index and a citation number. However, even the best in Engineering should publish 2 times a year or more papers to receive the same h-index as top scientists in overall domains.

The h-index distribution of domestic Engineering sub-disciplines has been presented and analysed in statistic categories. The suitable h-histograms and the cumulative probability density function (CPDF) have been elaborated for 21 sub-disciplines and thereupon the Engineering sub-disciplines have been arranged into three clusters. It has been demonstrated that Engineering as the whole and Engineering sub-disciplines are underestimated, compared to other academic disciplines. The adequate normalisation factors have been suggested.

Several other conclusions considered the h,H-indices as the measure of the knowledge progress addressed to individual researchers and to collective, e.g., journals, institutions, organisations, countries, adequately have been written. The h,H-indices are the general measure of the position of the given subject (person or organisation) but cannot be universal.

**Key words:** bibliometric indicators, citations, citation metrics, h-index, H-index, country rank, SCImago indicator, disciplines, domestic Engineering sub-disciplines, statistics.

## 1. Introduction

The paper aims at defining new challenges to the engineering community which result from the research policy and dominant tendency of parameterization of everything. Of course, it is the subjective and preliminary opinion. Mainly, it concerns effects of assessments of people and groups by taking into consideration their bibliometric measures, especially the Hirsch Index. The "parameterized" evaluation has features of objectivity and can be a useful tool while preparing ranking lists. However, it is necessary to mention that the evaluation applies to groups of researchers consisting of outstanding personalities, e.g., in the discipline civil engineering the evaluation concerns groups which deal with designing, building up, maintenance, repairing, demolition and lastly recycle the building structures. Those activities, consume huge amounts of matter (mass and energy); it is 40% of the overall world consumption. The durability/working life of the building objects, can be measured by multiplicity of a designer's life expectancy. Specific obligations are imposed on the creators, moreover, for that reason the "conservative" actions can be taken. This can be reflected in their work and determine the discipline [1]. Compared to the computer estimations, results are verified by experience and based on test results. Exactly the same rule for the parameterized evaluations of people can be used.

Interestingly, nobody asks a question why we should classify people by just a single number. It seems that there exists an inexpressible axiom that "course is inescapable" [2]. It is part of the general civilization madness enumerating everything. More than 267 million pages [3] on unstructured research data are available on line and thereby create dizzying sense of the knowledge infinity [4]. The ranking list of researchers could make sense of order in appearing chaos. There is the measure for measures [5]. "Trust in numbers" is treated as the synonymous of objectivity in science [6]. The point is even more delicate in the Engineering Science area where science and technology are interpenetrated. Once numbers are received, the conclusion often could be elaborated by mechanical methods, that is usually done by computer. However, it should be refreshed that numbers should firstly serve as a strategy of communication. K. Życzkowski statement [7]

*e-mail: l.czarnecki@il.pw.edu.pl

should be shared *bibliometric indices can be further developed but it is difficult to claim that exists a single number capable to quantity scientific achievement in an ambigious way*. However, it only concerns the actual value not the virtual space. A. Sawicki [8] said that only comparisons of similar with similar could be made. It is a far-reaching proposal, however, surely for the assessment and during the creation of ranking lists it is necessary to select, in a proper way, the "scale and tools". Paradoxically, parameterization is a common and well accepted way in the engineering methodology. The same method used, ad extremum, to evaluate engineering researches seems to provide erroneous conclusions.

**Conditions.** Considerations have been elaborated at two levels:

- careful thoughts of domestic experiences,
- universality of knowledge which implies permanent progress and the particularity of h-index as the potential measure of knowledge progress.

The parameterization of science and scientific researchers is the over-world tendency [9]. However, the direct inspiration for this study has been taken from domestic experiences. The Committee for Evaluation of Scientific Units is a consultative and advisory body to the Minister of Science and Higher Education. The main task of the Committee is to draw up the project of parameters and criteria for comprehensive evaluation of scientific units. Moreover, the Committee indicates to the Minister the leading scientific units taking into account the quality of their scientific activity in order to determine the level of financial support granted to fund their research potential. The National Science Centre is a government executive agency set up to fund basic research. Basic research is original experimental or theoretical research work undertaken primarily to acquire new knowledge of the underlying foundations of phenomena and observable facts, without the direct practical application or use. KEJN and NCN in most of the announced grants and projects have special criteria for the applicants. Namely, the applicants must give the number of all publications and H-index[1]. The division into a number of publications and a number of citations is important. The expectations that the publications will or should be noticed, especially by the browsers, are high. There are no additional limitations and comments about the H-index and the results not only will be important during the parameterisation of scientific units but also in the process of acquiring grants. The above tendency has been presented in the regulation of the Minister of Science and Higher Education which concerns the criteria, conditions and procedures of applying for obtaining the status of the Leading National Research Centre. There the following is said:

- when assessing applications in science, in the areas of technical science (...) the following is taken into account (. . . ) the number of citations and H-index estimated from the works published in journals, which are on the Journal Citation Report list, prepared for every research worker or university teacher,
- the number of published works has been limited to 25% of the best journals taken from the list of Journal Citation Report for the restricted areas of science and prepared by the Thomson's Reuters Scientific.

There are some similarities between a journal impact factor (JIF) and h-index. Particularly that the term "impact factor" has gradually evolved into description both journal and author impact. The journal impact factor is the average number of citations for an article in a particular journal during a given period of time (usually the last two years, sometimes previous five years). It has been found [10, 11] that between JIF and h-index there exists a strong and significant relation, which is not a surprise. The creation of h-index allowed to discern scientists who really did science that had an impact and JIF became a tool that underlines in which journal that science make an appearance. Now, after more than 50 years of JIF using, still arises the fundamental question: what is the JIF? If the JIF is not an outdated artifact rather than a steppingstone to journal certification [12]. The answer (Fig. 1) seems to be very close to the first option [12, 13]. Moreover, E. Garfield, 1955 [14], the creator of the JIF published paper [12, 13] entitled "The Agony and the Ecstasy – the History and Meaning of the Journal Impact Factor". Simultaneously, E. Garfield [14] by himself considered that more adequate to the situation would be another title: "Citation Sanity and Insanity – the Obsession and Paranoia of Citations and Impact Factors" [15].

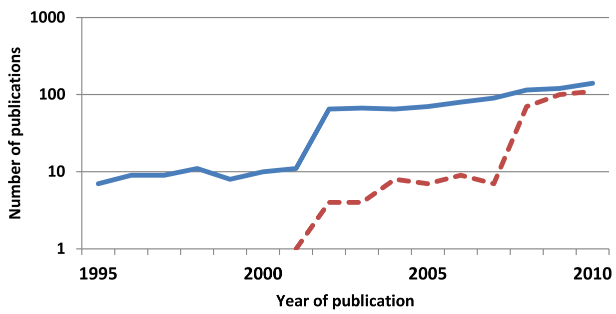



Fig. 1. Exponential increase in documents according to J.K. Vanclay (after Ref. 12) found with Scopus search for "journal impact factor" showing all documents (continuous line) and critical documents (dotted line) with words: "bias", "limitation", "problem", "manipulate", "misuse" or "flaw" placed in the abstract

On the margin of JIF discussions three groups of journals have been categorized [13]:

- high quality journal – published breakthroughs, new proposals and new types of applications,
- standard journal – published standard, yet interesting applications of known techniques,
- ordinary journal – published really routine applications which basically have only an archiving function.

[1]h or H – see Sec. 3 and the first conclusion. In the Ministry documents the capital letter H is used.

This categorization could be a contribution of interest to the trial of evolution of the state of Polish Engineering Science [16]. A. Osielawska [17] statistically proved that the fact of low contribution of Polish authors, dealing with Economy, into "international" journals does not mean that this community of researchers is limited only to the Polish science circle. The conclusion has been done on the base: which journals are mainly cited in the papers published in Polish journals. It is the most likely that the same statement could be done towards Engineering. One of the fundamental results of the JIF controversy is expressing a doubt if the science really progresses by measuring the impact of journals.

R. Rousseau [13] states the reason, that an impact factor is measured, as not scientific but as underlying cause which simply is involved with a human nature. If we try to transform the question into the h-index domain the situation will be even more delicate. After 2 years, since J. Hirsch declared his index J. Kuan-Teh, USA (h = 39), Editor-in-Chief of Retrovirology (JIF = 6.47) in the title of his paper [18] has put a question: *is it time to individualize citation metrics?* As his reply could be the sentence: *like it or not, use it or not, each author's personalized citation number and h-index are there for all to compare.* It means come back to a human case. The perception of a personal prestige is even the more murky subject than the perception of the journal prestige. It will be a pity if the function of h-index diminishes mainly to a visibility. As aforementioned in the Introduction we need the "scale and tool". For this reason we should distinguish an "artifact" [12] and a "tool". An "artifact" becomes a tool when knowledge and understanding about the proper use of the artifact develops in a person's mind [13]. On the following pages we try to answer to what extend the h-index is understood as a tool for research evaluation, particularly, in the engineering area. Perhaps it is necessary to get rid of some illusions of the illusionless territory. In this context Albert Einstein statement *Not everything that can be counted counts and not everything that counts can be counted* is sometimes remembered [19]. The great dramatic words can also be quoted *Are all thy conquests, glories, triumphs, spoils, shrunk to this little measure* (W. Shakespeare, Julius Caesar). The past experiences are not very promising. "Lost in publication", "measurement harms science" [20], "abuses" [13], "rank injustice" [21] are the phrase from titles and abstracts of relevant papers. The fundamental questions are [19]:

- does our academic system of activity rewarding, that addresses the question, matter most to the society?
- how to promote the advancement of knowledge if real power comes from great ideas?
- how to recognize individuals and institutions that best generate new knowledge?

## 2. h-nature of things

Jorge E. Hirsch is an Argentine American professor of physics at the University of California, San Diego. He published in 2005 the article 'An Index to Quantify an Individual's Scientific Research Output' in the Proceedings of the National Academy of Science [22]. Hirsch wrote: *I propose the index h, defined as the number of papers with citation number higher or equal to h, as a useful Index to characterise the scientific output of a researcher.*

Such proposition was quickly accepted. The article which was cited more than 2 000 times increases the author h-index (h = 25!) of only one point. It well reflects the significance of h-index and discloses one of the index peculiarities: how difficult it is to increase the index. Thus, the $h$-index reflects both the number of publications and the number of citations per publication. It can be expected that an author who has $N$-publications has certain $h$-index and $h$ publications have been at least cited h times and $N$-$h$ has been less cited or at all (Fig. 2). Although J.E. Hirsch created small $h$, in the NCN and KEJN documents capital letter $H$ appears. R. Kierzek [23] in his proposal used $h_m = \Sigma h/N^{0.4}$ where $N$ – stands for the number of publications in one institution. It seems to be rational to accept small $h$ for an individual output and capital letter $H$, the collective indicator, e.g. for organisations, institutions and countries, $H = h_m = \Sigma h/N^{0.4}$. It should be stressed that all non-productive publications decrease the H-index.

J. E. Hirsch measures related to the Nobel Prize laureates in Physics from 1985 to 2005 (Fig. 3). The author proved that the h-index would present productivity-quality-visibility. So it was. People, however, would have some questions to the visibility. Nowadays, in a professional CV prepared by a scientist $h$-index is indicated. It can be expected that in future $h$ will be placed on a visiting card, near the e-mail address. Additionally, in 2007 Hirsch added some prediction values to the $h$-index [24].

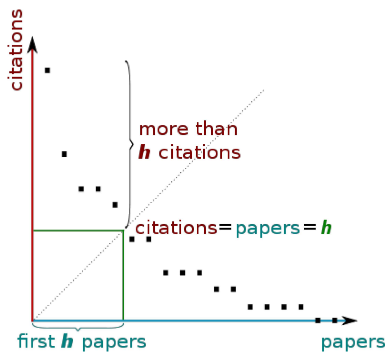



Fig. 2. H-index from a plot of decreasing citations for numbered papers (after Ref. 22)

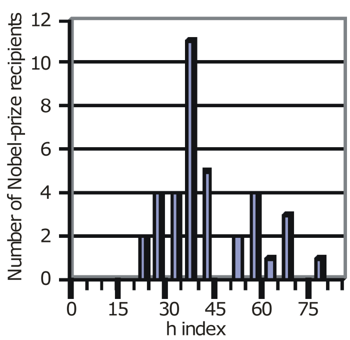



Fig. 3. The Nobel Prize laureates in physics (after Ref. 22)

The h-index can be measured with the use of different citation databases, e.g.:

- ISI Web Knowledge/Web of Science, Thomson-Reuters,
- SCOPUS, Elsevier,
- Publish or Perish (PoP), Google Scholar,
- Academic Search Microsoft.

There is a well known example that one of the most cited computer scientists, Hector Garcia-Molina ($h = 120$), Stanford, USA has gathered nearly 60 000 citations in PoP, but in ISI has received just a few hundred citations [19]. It is due to the fact that the most of his papers have been published and cited in conference proceedings which are appreciated by Google Scholar but they were not recognized by Thomson Reuters ISI. The PoP browser seems to be the most favourable to engineering disciplines. Moreover, this database frequently recognises books not only papers in journals; books and articles written in Polish are also in its resources. It should be claimed that $h$-information always goes together with an adequate citation database. However, the author should have a right to free selection of a suitable database yet, under the condition that unnoticed citation does not mean non-existing. Even so, it can be assumed that PoP finds only 1/5 publications and 1/3 citations of Polish scientists (see Sec. 6). The concept of H-index gained huge popularity, appreciation and, contributed to the development of scientometrics, the science of knowledge measurements.

The H-index can also be created for countries. Poland is on the $24^{th}$ position with H = 2 81, in comparison USA = 1305 (Table 1). Poland ranks $33^{rd}$ in classification by population whereas by Gross Domestic Product, GDP the country holds the $20^{th}$ position. In the area of 'Engineering', Poland is on the $30^{th}$ position with H = 82 where USA is still on the top with H = 485 (Table 2).

Table 1
Country Ranking according to H-index, all areas 1996–2011: SJR (December $5^{th}$, 2012) (after Ref. 25)

| | Country | Documents | Citable documents | Citations | Self-Citations | Citations per Document | H index |
|---|---|---|---|---|---|---|---|
| 1. | United States | 6 149 455 | 5 738 593 | 114 546 415 | 54 226 872 | 20.51 | 1 305 |
| 2. | United Kingdom | 1 711 878 | 1 550 373 | 27 919 060 | 6 703 673 | 18.03 | 802 |
| 3. | Germany | 1 581 429 | 1 490 140 | 23 229 085 | 6 171 727 | 16.19 | 704 |
| 4. | France | 1 141 005 | 1 073 718 | 16 068 688 | 3 749 874 | 15.58 | 646 |
| 5. | Canada | 885 197 | 836 836 | 13 928 114 | 2 727 913 | 18.19 | 621 |
| 6. | Japan | 1 604 017 | 1 563 732 | 18 441 796 | 5 520 032 | 12.09 | 602 |
| 7. | Italy | 851 692 | 803 004 | 11 279 167 | 2 639 721 | 15 | 550 |
| 8. | Netherlands | 487 784 | 457 933 | 8 928 850 | 1 524 755 | 20.82 | 545 |
| 9. | Switzerland | 350 253 | 329 198 | 6 873 551 | 966 536 | 22.46 | 537 |
| 10. | Sweden | 337 135 | 321 725 | 6 111 804 | 1 005 775 | 19.78 | 484 |
| 11. | Australia | 592 533 | 551 667 | 8 180 664 | 1 770 774 | 16.65 | 481 |
| 12. | Spain | 665 977 | 623 236 | 7 640 544 | 1 958 835 | 13.66 | 448 |
| 13. | Belgium | 265 913 | 251 632 | 4 161 308 | 630 041 | 17.81 | 428 |
| 14. | Denmark | 183 880 | 173 771 | 3 444 509 | 514 632 | 21.17 | 399 |
| 15. | Israel | 204 262 | 194 752 | 3 283 119 | 483 335 | 17.35 | 393 |
| 16. | Austria | 188 440 | 177 324 | 2 688 324 | 387 884 | 16.51 | 355 |
| 17. | China | 2 248 278 | 2 226 529 | 9 288 789 | 5 014 506 | 6 | 353 |
| 18. | Finland | 170 476 | 165 195 | 2 771 982 | 462 377 | 18.28 | 352 |
| 19. | South Korea | 497 681 | 487 459 | 3 988 716 | 917 147 | 10.32 | 309 |
| 20. | Norway | 141 143 | 133 311 | 2 021 938 | 339 172 | 17.19 | 308 |
| 21. | Russian Federation | 527 442 | 521 993 | 2 811 862 | 837 763 | 5.49 | 308 |
| 22. | Brazil | 391 589 | 378 540 | 2 884 793 | 965 615 | 9.96 | 285 |
| 23. | India | 634 472 | 602 868 | 3 860 494 | 1 335 686 | 7.71 | 281 |
| **24.** | **Poland** | **304 003** | **297 361** | **2 149 143** | **571 333** | **8.13** | **281** |
| 25. | Hong Kong | 144 935 | 139 331 | 1 722 546 | 262 368 | 13.52 | 268 |
| 26. | New Zealand | 114 495 | 107 441 | 1 504 946 | 248 529 | 15.43 | 264 |
| 27. | Ireland | 91 125 | 85 341 | 1 149 729 | 141 683 | 16.18 | 254 |
| 28. | Taiwan | 351 610 | 343 223 | 2 825 736 | 696 835 | 10.08 | 249 |
| 29. | Greece | 160 760 | 152 000 | 1 589 963 | 289 460 | 11.93 | 247 |
| 30. | Singapore | 126 881 | 122 436 | 1 330 684 | 191 033 | 12.51 | 240 |
| 31. | Hungary | 100 137 | 96 842 | 1 058 391 | 182 169 | 11.57 | 239 |
| 32. | Czech Republic | 142 090 | 137 882 | 1 103 719 | 272 685 | 9.14 | 223 |
| 33. | Portugal | 117 469 | 113 411 | 1 150 280 | 234 405 | 12.77 | 218 |
| 34. | Mexico | 144 997 | 140 713 | 1 174 802 | 259 075 | 9.83 | 216 |
| 35. | South Africa | 107 976 | 101 434 | 1 013 102 | 225 507 | 11.11 | 216 |

Table 2
Country Ranking according to H-index; Engineering area 1996–2011: SJR (December $5^{th}$, 2012) (after Ref. 25)

| | Country | Documents | Citable documents | Citations | Self-Citations | Citations per Document | H index |
|---|---|---|---|---|---|---|---|
| 1. | United States | 646 693 | 628 373 | 6 054 196 | 2 335 391 | 9.68 | 485 |
| 2. | United Kingdom | 142 413 | 138 050 | 1 179 649 | 280 089 | 8.65 | 242 |
| 3. | Germany | 136 590 | 133 452 | 892 300 | 224 568 | 7.05 | 218 |
| 4. | Canada | 95 254 | 93 085 | 753 251 | 158 979 | 8.7 | 194 |
| 5. | France | 101 268 | 99 498 | 772 071 | 212 178 | 8.45 | 193 |
| 6. | Japan | 198 677 | 196 307 | 1 015 526 | 346 698 | 5.35 | 193 |
| 7. | Italy | 78 870 | 77 069 | 607 837 | 158 080 | 8.59 | 173 |
| 8. | China | 506 147 | 503 812 | 1 235 737 | 823 142 | 3.23 | 172 |
| 9. | Switzerland | 26 898 | 26 208 | 292 015 | 42 731 | 12.02 | 167 |
| 10. | Netherlands | 36 915 | 35 983 | 364 479 | 64 523 | 10.94 | 163 |
| 11. | Australia | 44 571 | 43 368 | 371 136 | 73 709 | 10.08 | 153 |
| 12. | South Korea | 90 685 | 89 555 | 513 995 | 128 489 | 6.62 | 141 |
| 13. | Spain | 50 243 | 49 276 | 353 652 | 100 834 | 8.59 | 140 |
| 14. | Hong Kong | 29 610 | 29 065 | 283 544 | 46 871 | 10.01 | 140 |
| 15. | Singapore | 29 567 | 29 098 | 270 686 | 43 233 | 9.65 | 136 |
| 16. | Taiwan | 76 165 | 75 212 | 480 568 | 150 780 | 7.24 | 136 |
| 17. | Sweden | 28 956 | 27 553 | 239 100 | 42 815 | 8.67 | 134 |
| 18. | Belgium | 22 273 | 21 845 | 203 861 | 37 026 | 10.15 | 129 |
| 19. | Israel | 17 375 | 17 069 | 173 818 | 28 678 | 10.5 | 129 |
| 20. | India | 59 190 | 58 065 | 309 724 | 89 493 | 6.27 | 126 |
| 21. | Denmark | 11 333 | 11 053 | 118 625 | 18 151 | 11.84 | 112 |
| 22. | Finland | 16 240 | 15 992 | 120 949 | 23 178 | 8.03 | 104 |
| 23. | Brazil | 26 297 | 25 902 | 133 008 | 34 172 | 6.09 | 98 |
| 24. | Greece | 18 078 | 17 658 | 130 342 | 26 629 | 8.16 | 97 |
| 25. | Turkey | 22 541 | 22 074 | 152 590 | 41 232 | 8.01 | 92 |
| 26. | Russian Federation | 64 657 | 64 164 | 138 659 | 51 133 | 2.14 | 91 |
| 27. | Norway | 10 200 | 9 992 | 75 834 | 12 759 | 8.85 | 90 |
| 28. | Austria | 14 491 | 14 092 | 87 562 | 16 802 | 7.14 | 87 |
| 29. | Portugal | 13 414 | 13 071 | 90 178 | 21 194 | 8.19 | 84 |
| 30. | **Poland** | **28 669** | **28 286** | **102 202** | **32 926** | **4.13** | **82** |
| 31. | Ireland | 8 523 | 8 346 | 57 152 | 10 158 | 7.69 | 78 |
| 32. | New Zealand | 6 626 | 6 455 | 49 934 | 8 244 | 8.6 | 77 |
| 33. | Mexico | 12 871 | 12 672 | 58 819 | 15 213 | 5.39 | 76 |
| 34. | Iran | 21 853 | 21 484 | 86 710 | 32 965 | 6.85 | 71 |
| 35. | Czech Republic | 9 421 | 9 280 | 46 369 | 11 975 | 6.11 | 70 |

Some categories[2] of Engineering remain close in rank: Biomedical Engineering (33), Civil & Structural Engineering (32), Electrical & Electronic Engineering (32), Mechanical Engineering (30). Some categories are slightly better ranked: Ocean Engineering (28) and Safety Risk, Reliability and Quality (25). The Architecture falls significantly behind (69!). In the recent 15 years the number of papers published by the Polish authors in all areas has increased almost 3 times including the Engineering area as well (Fig. 4a,b). Over the same period, the relevant world increase is 2 times bigger. A published article should have enough time to be cited. The period of 7 years since the publication date seems to be needed (Fig. 5). However, the papers which achieved a high impact factor are believed to be usually cited within months since publication and certainly within a year or so [14]. The relation between cited and uncited Polish papers after more than 7 years since publication has been established at the level of 20% uncited for all area and 40% (!) uncited for Engineering (Fig. 4c,d). Respective numbers for the world are around 25 and 30%; for USA 15 and 30%. All the data in this section have been taken out from SJR Scimago Journal & Country Rank, where they are published under meaningful motto from Horatio “Est modus in rebus” – “There are proper measures in things”.

There are several axiomatic characterizations [26] of the Hirsch index:

- Hirsch index reflects simultaneously the number of publications and the number of citations per publication,
- authors with very few high-impact publications and authors with very low-impact publications will score weak $h$-index,
- the $h$-index of a given scientist never decrease on timeline,

[2]Categorisation according to Scientific Journal Ranking

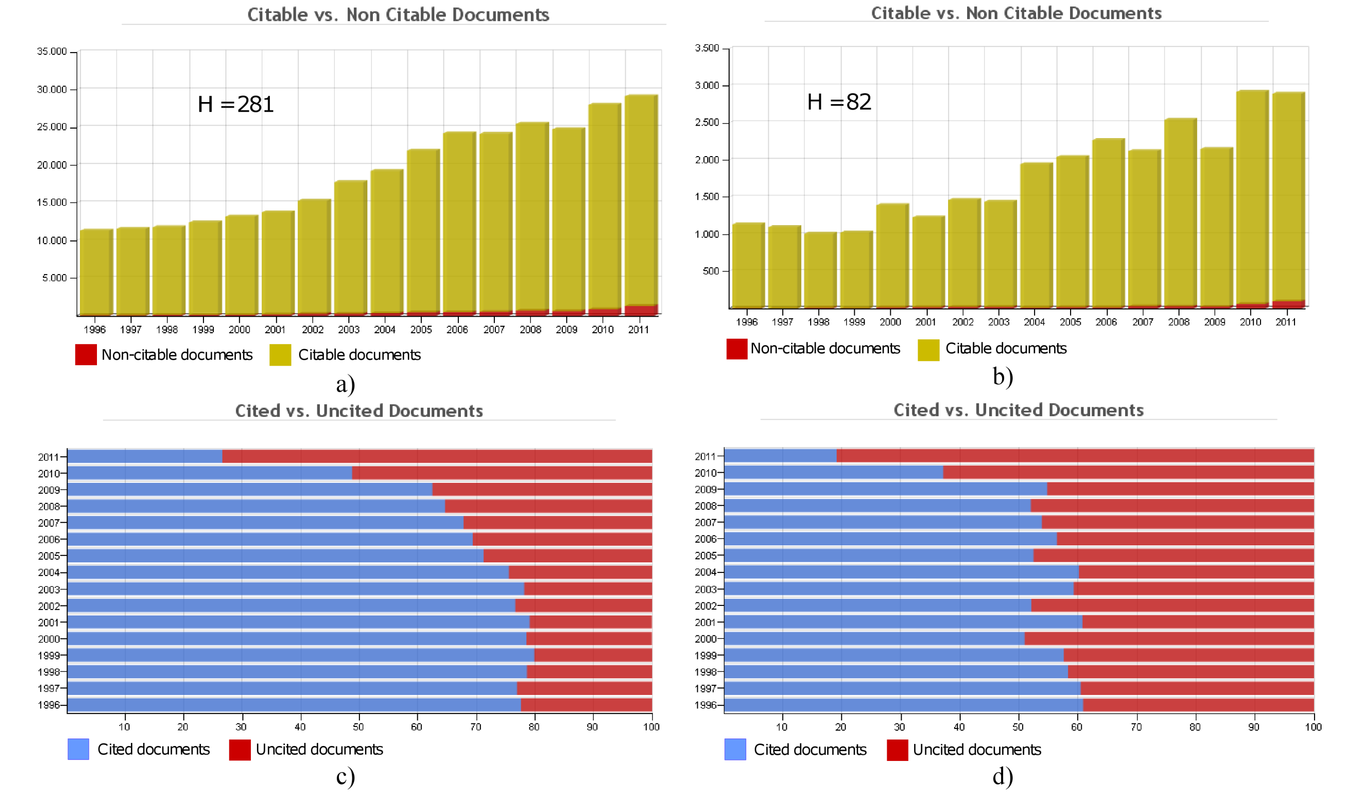



Fig. 4. Number of papers of Polish authors (a, b) in all areas (a) and in Engineering (b) and relation of cited and uncited papers (c, d) in all areas (c) and in Engineering (d) (after Ref. 25)

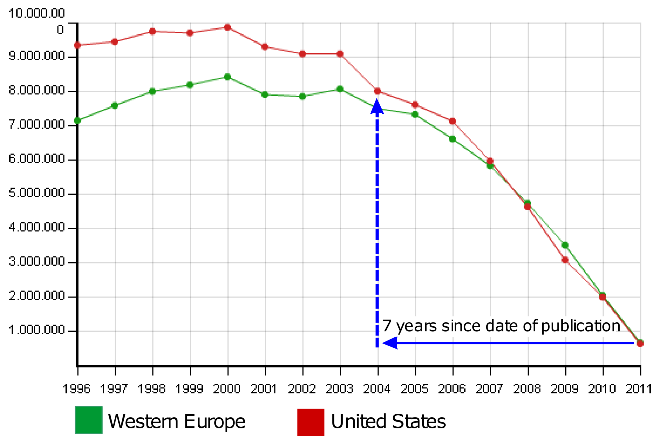



Fig. 5. Citations number in the USA and Western Europe; the justification taken out in 7 years period of time for data stabilization (after Ref. 25)

- the $h$-index of a given scientist may increase when a new paper is published and will attract citations significantly, vis. above current $h$-level,
- if a new publication is only "average" with respect to the current index it should not raise the index; publications cited less frequently than $h$-times are "wasted",
- if an existing paper attracts new citations it should raise the index only if the number of citations per paper increase above the current index value (Fig. 6),

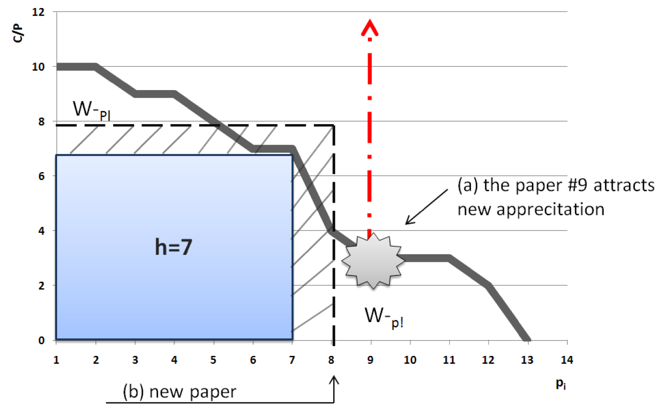



Fig. 6. Citations per papers (C/P) vs. paper numbers ($p_i$) ordered according to decreasing citations (after Ref. 26). The area under curve gives the total number of citations. The square yields suit $h$-index. W-$_{Pl}$ -waste top area, potential less – even significant increment of citations change nothing in $h$-index. W-$_{p!}$ – waste area at the bottom but still with a potential. The ways of $h$-index increase until $h$+1: a) existing paper, e.g. number 9, attracts new readers and received citations above the current $h$-value, b) new paper gained more than $h$ citations and will be located on the new position $p_i = 8$, if the number of citation is higher than current $h$ (C/P>h)

- $h$-index is an integer number by increasing the number of citations to a single article, the $h$-index should not increase by more than 1; the "extra top publication" cited more (even much more) than $h$-times are "wasted",
- the $h$-index is absolutely neutral to the context of citations: affirmative or negative,

- a research institution has H-index [27,28] of $i$ when at least $i$ researchers from that institution have $h$-index of at least $i$, (it is not the same as R. Kierzek proposal [23]),
- the co-authorship effect, could be calculated [29], by dividing $h$ by the mean number <Na> of authors (author multiple occurrences are allowed) in the considered $h$ papers: <Na>= Na$^{(T)}$/h where Na$^{(T)}$ – is the total number of authors. Thus, we obtain a new index $h_I$: $h_I = h/$<Na>$=$ $h^2$/Na$^{(T)}$. A research with index $h_I$ has $h_I$ collective author papers with at least $h_I$ citation if he/she had published alone. If a given researcher is the only author of his/her $h$ papers, than Na$^{(T)}$ $= h$ and $h_I = h$ in this case. The justification for this procedure is that more authors could produce more future self-citations which in turn could affect more readers and finally may produce statistical biases. Till now it is just a suggestion of P.D. Batista et al. [29], which seems to be rational, not yet accepted in common.

## 3. Shaping the H-index in various Engineering sub-disciplines

Despite all attempts to refine [26, 30], adjust [2, 31] or improve [28, 32, 33] $h$-index; Hirsch's original measure prevailed and has already become a standard tool in evaluative bibliometric [29]. On the other hand, there is an unquestionable opinion that citation of Engineering science is much lower than, e.g., of a majority of Physics and Chemistry [16, 29]. From an extensive and statistically proved studies made by E. Lillquist et al. [30] someone could learn that the selected sciences could be categorized as in Fig. 7 according to the decreasing median of $h$-index value, as follow: Physics, Biology, Chemistry, Chemistry Engineering, Electrical, Mechanical, Math, Civil Engineering. The limit values are: 32 and 5. J. Hirsch suggests that the $h = 12$ is the value adequate for "matured" scientists. In this context, the situation of Engineering and its sub-disciplines do not look very nice. Comparison of the publication and citation profiles as cumulative probability density functions (CPDF) (Fig. 8) shows a shift in a greater number, more citations than publications by factor more than 10 but the order of disciplines are basically the same (compare Fig. 7).

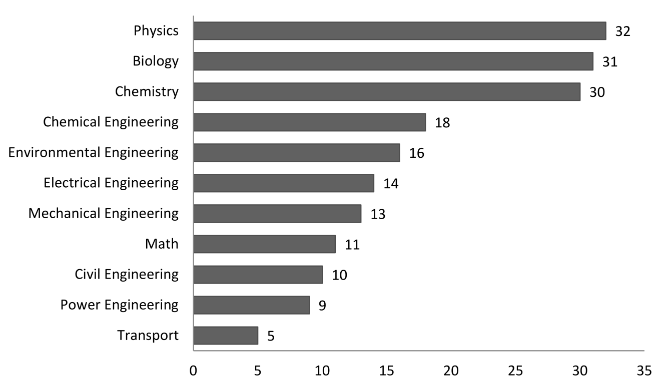



Fig. 7. The median – H-index value for full professors community in various sciences (after Ref. 30)

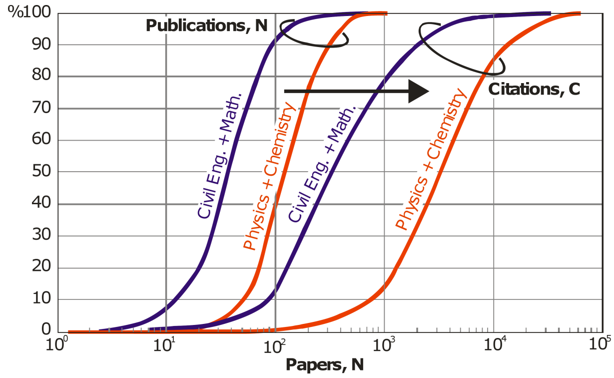



Fig. 8. The profile of publication and citation for full professors (Ref. 30); the field of outline covers disciplines: mathematics and civil engineering (the same line), mechanics, electrical engineering, chemical engineering, biology, chemistry, physics are in the same order in both cases.

Various attempts have been done to scaling or normalizing of $h$-index for different scientific fields [2]. The situation is even more complicated due to the fact that $h$-index represents simultaneously influence of total number of papers, $N$, and the rate of citation, $R : h = f(N, R)$. The Glänzel's model [34, 35] connects the $h$-index with those two fundamental scientometric indicators:

$$h = k \cdot N^{1/3} \cdot R^{2/3}. \quad (1)$$

On this base E. Csajbok et al. [36] elaborated the master curve (Fig. 9) for 20 disciplines of 40 countries. Taking this opportunity, authors made several points on the country of origin effects:

- when the $h$-index combines effort of size (number of papers) with that of impact (citation rate), smaller countries have not got much chance to compete with giants,
- EU countries have a strong positions in each field but none of them can successfully compete with the USA,
- newly accessed countries, like Poland, which belongs to large countries, is permanently ranked bellow smaller EU-55 countries like Belgium, Denmark or Austria.

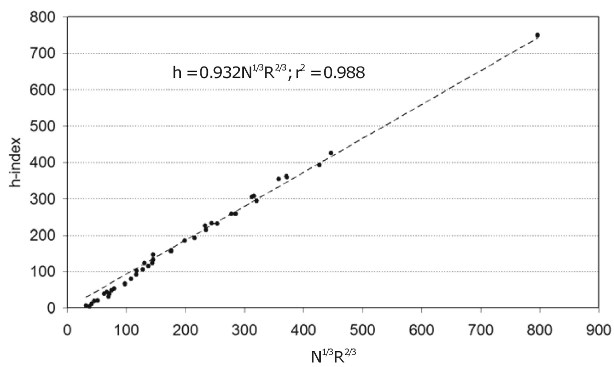



Fig. 9. Master curve $h = f(N, R)$ (after Ref. 36)

J. Podlubny [31] has found, that in the given period of time, the ratio of citations number, $C$, in any two fields of science remains close to constant, e.g.

$$C_{Ph}/C_E = 4, \quad (2)$$

where $C_{Ph}$ – number of citations for Physics and $C_E$ – number of citations for Mathematics.

Using the Glänzel's model (1) verified by E. Csajbok et al. [36] it could be calculated that for the same number of publications $N_{Ph} = N_E = N$ the relation between $h$-indices for Physics and Engineering is like below:

$$h_{Ph} = 2.5 h_E. \tag{3}$$

From the same relation comes that to receive the same value of $h$ for scientists in engineering he or she should publish two times more papers than scientists in physics.

$$N_{Ph} = 2 N_E. \tag{4}$$

More general scaling of the $h$-index has been proposed by J.E. Iglesias and C. Pecharroman [22]. Following the extend study of h-index distribution they proposed the normalized $h_N$:

$$h_N = f \cdot h, \tag{5}$$

$$h_N = 1.7h. \tag{6}$$

Table 3
The normalizing factor f: $h_N = f \cdot h$

| Discipline | $f$ |
|---|---|
| Agricultural Sciences | 1.27 |
| Biology & Biochemistry | 0.60 |
| Chemistry | 0.92 |
| Clinical Medicine | 0.76 |
| Computer Science | 1.75 |
| Economic & Business | 1.32 |
| Engineering | 1.70 |
| Environment/Ecology | 0.88 |
| Immunology | 0.52 |
| Materials Science | 1.36 |
| Mathematics | 1.83 |
| Microbiology | 0.63 |
| Molecular Biology & Genetics | 0.44 |
| Neuroscience & Behaviour | 0.56 |
| Pharmacology & Toxicology | 0.84 |
| Physics | 1.00 |
| Plant & Animal Science | 1.08 |
| Psychiatry/Psychology | 0.88 |
| Social Sciences, general | 1.60 |
| Space Science | 0.74 |

## 4. Lesson from the best

It is the fact that there exists a different science citation custom [2] which is reflected in the significant statistical measures such as mean, mode and median which value is counted per paper. It is also demonstrated in [2] that in Physics, Chemistry or Biology there exist many researchers with $h > 10$, on the other hand, in mathematics dominates $h < 10$, similar to Civil Engineering [30]. Moreover, the $h$-index of the top scientists in various academic disciplines (Table 4) could differ even several times ($h_{\min} = 23$, $h_{\max} = 111$ (146)). However, it is valid not only on the overall level, e.g., in Engineering or other disciplines, but also among Engineering sub-disciplines ($h_{\min} = 12$, $h_{\max} = 53$) (Table 5).

Here (Tables 4 and 5) an unique set of "the best" 15 Top Scientists in various domains and 17 Top Scientists from Engineering has been gathered. Generally, even "the best" should work hard, which means more papers – higher index (Fig. 10):

- for overall disciplines:

$$h_o = k_o\sqrt{N} = 4.2\sqrt{N}; \quad r = 0.79, \tag{7}$$

- for Engineering sub-disciplines

$$h_e = k_e\sqrt{N} = 2.7\sqrt{N}; \quad r = 0.83 \tag{8}$$

where $h_o$, $h_e$ – Hirsch index for overall disciplines and for Engineering sub-disciplines (adequately), $N$ – number of publications, $k_o$ and $k_e$ – constants of proportionality, and $r$ – correlation coefficient.

However, to receive the same $h$-index, engineering sciences should publish even 2 times more papers $\left[(\mathrm{ko}/\mathrm{ke})^2 = 2.4\right]$ than in overall domains. There exists a factor which differentiates an individual author's effort in the given domain. It is the citation rate $R_c = C/N$ – a number of citations per paper. Coefficients $R_c$ are significantly different ($R^o_{c,\min} = 18$, $R^o_{c,\max} = 282$) for various disciplines, moreover, the same occurs taking into consideration Engineering (Table 5) sub-disciplines ($R^E_{c,\min} = 13$, $R^E_{c,\max} = 133$). It could be estimated that "the best" who receive $h = 40$ should publish more than 200 papers in Engineering but in overall disciplines 100 papers appears to be sufficient. They should work hard but in Engineering harder.

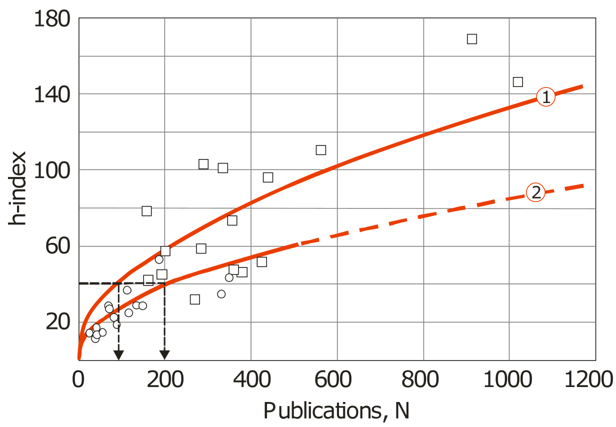



Fig. 10. The $h$-index vs. number of publications, $N$ for top scientists in various (see Table 4) academic disciplines (1) and for Engineering (see Table 5) sub-disciplines (2). It shows how many papers should be published to receive $h = 40$

Table 4
The $h$-index ranking of top scientists in various academic disciplines: Academic Search Microsoft, January $9^{th}$, 2013 (after Ref. 37)

| No. | Academia discipline* | Top author | Organisation | $h$-index** | Number of publication, N | Number of citations, C | Average Rate of citation per paper, R=C/N |
|---|---|---|---|---|---|---|---|
| 1. | Art and Humanities | Robert E.M. Hedges | University of Oxford | 23 | 85 | 2 907 | 34.2 |
| 2. | Agricultural Science | Patrick F. Fox | University College Cork National University of Ireland | 32 | 271 | 4 921 | 18.3 |
| 3. | Social Science | Douglas S. Massey | Princeton University | 42 | 163 | 8 747 | 52.0 |
| 4. | Mathematics | Pierre-Louis Lions | Paris Dauphine University | 45 | 194 | 12 378 | 68.8 |
| 5. | Environmental Sciences | Kevin C. Jones | Lancaster University | 46 | 379 | 8 632 | 22.8 |
| 6. | Material Science | Anthony G. Evans | University of California Santa Barbara | 47 | 361 | 11 385 | 32.0 |
| 7. | Multidisciplinary | Herbert Simon | Carnegie Mellon University | 52 | 424 | 52 978 | 125 |
| 8. | Chemistry | Janet Thornton | European Bioinformatics Institute EMBL | 57 | 202 | 32 017 | 158.5 |
| 8A | | [G. M. Whitesides] | Harvard University | [169] | [910] | | |
| 9. | Engineering | Ted Belytschko | Northwestern University | 58 | 283 | 18 284 | 64,6 |
| 10. | Geosciences | Yoram Kaufman | National Aeronautics and Space Administration, USA | 74 | 356 | 20 168 | 56.7 |
| 11. | Economics and Business | Andrei Ikjkl Shleifer | Harvard University | 78 | 159 | 44 758 | 281.5 |
| 12. | Computer Science | Scott J. Shenker | University of California Berkley | 96 | 440 | 45 621 | 103.7 |
| 12A | | [Herbert A. Simon] | Carnegie Mellon University | [146] | [1018] | | |
| 13.<br>13A | Physics | Edward Witten | Institute for Advanced Study | 101<br>[120] | 337<br>[398] | 41 013 | 121.7 |
| 14. | Biology | Eric S. Lander | Massachusetts Institute of Technology | 103 | 292 | 10 638 | 36.4 |
| 15. | Medicine | Karl Friston | University College London | 111 | 561 | 57 434 | 196.7 |

* Categorisation according to Microsoft Academic Search

** The Microsoft Academic Search has indexed 38 million publications and 18 million authors. H-index values may differ from other search engines. In a few cases in square brackets someone can find data from other sources: 8A [38], 12A [39], 13A [40]

Table 5
The $h$-index ranking of top scientists in various sub-disciplines of Engineering: Microsoft Academic Search, January $9^{th}$, 2013 (after Ref. 37)

| No. | Engineering sub-discipline* | Top author | Organisation | $h$-index | Number of publication, N | Number of citations, C | Average Rate of citation per paper, R=C/N |
|---|---|---|---|---|---|---|---|
| 1. | Mining Engineering | Lanru Jing | Royal Institute of Technology | 12 | 39 | 689 | 17.7 |
| 2. | Construction | H. Randolph Thomas | Pennsylvania State University | 14 | 44 | 567 | 12.9 |
| 3. | Nuclear Engineering | G. C. Sih | East China University of Science and Technology | 15 | 56 | 787 | 14.1 |
| 4. | Nanotechnology | Yusuf Altintas | University of British Columbia | 15 | 26 | 3 022 | 116.2 |
| 5. | Ocean Engineering | James Kirby | University of Delaware | 17 | 41 | 3 003 | 73.2 |
| 6. | Reliability and Risk Analysis | Mihailo D. Trifunac | University of Southern California | 19 | 90 | 3 458 | 38.4 |
| 7. | Transportation Engineering | Chandra R. Bhat | University of Texas Austin | 24 | 79 | 3 050 | 38.6 |
| 8. | Aeronautics and Aerospace Engineering | Yaakov Bar-Shalom | University of Connecticut | 25 | 117 | 11 730 | 100.3 |
| 9. | Chemical Engineering | Ignacio E. Grossmann | Carnegie Mellon University | 26 | 75 | 7 780 | 103.7 |
| 10. | Manufacturing Technology | Petar V. Kokotovic | University of California Santa Barbara | 28 | 72 | 9 589 | 133.2 |
| 11. | Biomedical Engineering | Julian C. R. Hunt | University College London | 29 | 148 | 5 301 | 35.8 |
| 12. | Civil Engineering | Zdenek P. Bazant | Northwestern University | 29 | 139 | 3 544 | 25.5 |
| 13. | Energy | Roy Billinton | University of Saskatchewan | 35 | 331 | 7 200 | 21.8 |
| 14. | Industrial Engineering | Thomas A. Lipo | University of Wisconsin Madison | 37 | 113 | 10 768 | 95.3 |
| 15. | Electrical and Electornic | Tatsuo Itoh | University of California Los Angeles | 44 | 349 | 10 261 | 29.4 |
| 16. | Mechanical Engineering | Ted Belytschko | Northwestern University | 53 | 189 | 18 284 | 96.7 |
| 17 | Engineering – overall | Ted Belytschko | Northwestern University | 58 | 283 | 18 284 | 64.6 |

* Categorisation according to Microsoft Academic Search. It is not clear what the difference between Construction (2.) and Civil Engineering (12.) is.

To be a Top Scientists one should not only be talented and hard-working but also should be interested in his/her scientific progress in a degree enabling him/her to get results or at least work in the field with an adequate citation to attract readers to citation. J.E. Hirsch Fig. 11 [22] found that between $h$-index and a number of citations, $C$ is exponential relation

$$C = \alpha h^2, \tag{9}$$

where $\alpha$ is proportionally constant, empirically $\alpha$ = 3–5. It means:

$$h = k\sqrt{c} = \frac{1}{\sqrt{\alpha}} \cdot \sqrt{c}, \tag{10}$$

$$h_o = 0.42\sqrt{c}; \quad r = 0.62, \tag{11}$$

$$h_e = 0.38\sqrt{c}; \quad r = 0.89. \tag{12}$$

It means that $\alpha_o = 5.7$ and $\alpha_e = 6.9$. Probably, it is the biggest influence of "big hits". In such case J. Hirsch [22] implies untypical value $\alpha > 5$.

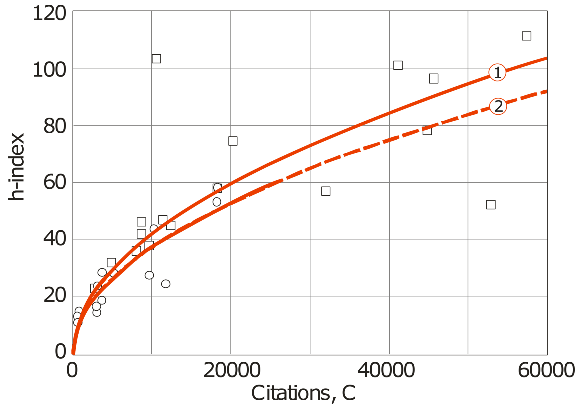



Fig. 11. The h-index vs. citation number for top scientists in various (see Table 4) academic disciplines (1) and for Engineering (see Table 5) sub-disciplines

## 5. The h-index distribution of domestic Engineering sub-disciplines

The Representative Set (RS) of scientists from various domestic Engineering sub-disciplines: Architecture and Urban Planning (A), Production Engineering (PE), Transport (T), Mining (M), Management of Mineral Resources (MM), Machine Building (MB), Civil Engineering & Hydroengineering (CEH), Geodesy (G), Metallurgy (Met), Metrology and Scientific Instrumentation (MSI), Electrical Engineering (E), Electronics & Telecommunication (E&T), Informatics (I), Automatic Control & Robotics (A&R), Biocybernetics & Biomedical Engineering (B&B), Acoustics (Ac), Environmental Engineering (EE), Thermodynamics & Combustion (TC), Chemical & Process Engineering (ChE), Mechanics (Mech), Materials Science (MS) has been selected. It is obvious (Table 6, Fig. 12a–x) that the histograms of $h$-index for RSs are sometimes very dispersed (E, E&T, I, Mech, MS, B&B, A&R, ChE). Often positive outliers are observed, usually with $h$-value several times larger than mode. In the case of Automatic Control and Robotics even 12 times. However, the common feature, it is sorry to say, is a relatively low median (2–12) and even lower mode (2–10). Following the verification of the significance of differences between RS variances (Table 7) and means (Table 8) the top cluster (E, E&T, I, A&R, B&B, ChE, Mech, MS), upper medium cluster (EE, Met, MSI, Ac, T&C) and lower medium cluster (T, M, MM, MB, C&H, G) can be recognised from "others" (A, PE) (Fig. 13). It allows to restructure statistics measures of the h-index distribution of domestic Engineering sub-disciplines (Table 9).

In many cases, if not in all, the domestic medians (Fig. 14) are significantly lower than the world median (compare Fig. 7). We can assume several reasons for that, which are valid generally for "domestic" output:

- country-of-origin effect,
- non-English authors,
- publishing (partially) in domestic journals.

However, some particularities concerning Engineering sub-disciplines should be considered. The very low median is an attribute of the discipline which "applied product" is addressed basically to local-domestic: A, T, M, C&H, G. It should be also stressed that behind some publications contribution to the lab-experimental works in many engineering sub-disciplines stands high.

Table 6
Statistic measure of $h$-index distribution of domestic Engineering sub-disciplines

| No. | Engineering sub-discipline | Observation number, n | H min | H max | H range | H mean | H median | H mode | Standard deviation $\sigma$ | Variation coeff. $\sigma/H$ | Skewnes coeff. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. | Architecture and Urban Planning, A | 29 | 1 | 5 | 4 | 2.07 | 2 | 2 | 1.13 | 0.55 | 1.13 |
| 2. | Production Engineering, PE | 22 | 0 | 9 | 9 | 2.82 | 2 | 2 | 2.28 | 0.81 | 1.54 |
| 3. | Transport, T | 30 | 0 | 10 | 10 | 3.07 | 3 | 3 | 1.80 | 0.59 | 1.95 |
| 4. | Mining, M | 39 | 0 | 12 | 12 | 3.30 | 3 | 2v4 | 2.72 | 0.82 | 1.49 |
| 5. | Management of Mineral Resources, MM | 30 | 0 | 9 | 9 | 3.53 | 3 | 3 | 2.03 | 0.57 | 0.82 |
| 6. | Machine Building, MB | 37 | 0 | 12 | 12 | 4.14 | 4 | 4 | 2.50 | 0.60 | 1.06 |
| 7. | Civil Engineering and Hydroengineering, CEH | 43 | 1 | 18 | 17 | 4.56 | 3 | 3 | 3.25 | 0.71 | 2.17 |
| 8. | Geodesy, G | 33 | 2 | 12 | 10 | 4.79 | 4 | 4 | 2.34 | 0.49 | 1.33 |
| 9. | Environmental Engineering, EE | 38 | 1 | 32 | 31 | 5.08 | 3 | 1v3 | 5.50 | 1.08 | 3.42 |
| 10. | Metallurgy, Met | 30 | 0 | 16 | 16 | 5.40 | 4 | 2 | 4.14 | 0.76 | 1.11 |
| 11. | Metrology and Scientific Instrumentation, MSI | 33 | 0 | 24 | 24 | 5.85 | 4 | 1v2 | 5.84 | 1.00 | 1.80 |
| 12. | Electrical Engineering, E | 46 | 1 | 28 | 27 | 5.89 | 5 | 2v3 | 5.69 | 0.97 | 2.41 |
| 13. | Acoustics, Ac | 31 | 1 | 17 | 16 | 5.97 | 6 | 6 | 4.14 | 0.69 | 1.05 |
| 14. | Thermodynamics & Combustion, TC | 28 | 1 | 19 | 18 | 6.39 | 6 | 5v6 | 3.92 | 0.61 | 1.85 |
| 15. | Electronics & Telecommunication, E&T | 46 | 0 | 29 | 29 | 7.13 | 6 | 4 | 6.33 | 0.89 | 1.63 |
| 16. | Informatics,I | 43 | 0 | 33 | 33 | 7.67 | 5 | 0v3 | 7.7 | 1 | 1.43 |
| 17. | Automatic Control and Robotics, A&R | 37 | 1 | 46 | 45 | 8.3 | 5 | 4 | 8.52 | 1.02 | 2.81 |
| 18. | Biocybernetics & Biomedical Engineering B&B | 32 | 2 | 38 | 36 | 9.97 | 8 | 4v10 | 7.35 | 0.74 | 2.06 |
| 19. | Chemical and Process Engineering, ChE | 29 | 3 | 27 | 24 | 10.14 | 9 | 4 | 5.95 | 0.59 | 0.89 |
| 20. | Mechanics, Mech | 44 | 3 | 36 | 33 | 11.66 | 10 | 6 | 7.33 | 0.63 | 1.33 |
| 20a | Mechanics, Mech (Scopus)* | 44 | 1 | 23 | 22 | 7.07 | 6 | 4 | 5.19 | 0.73 | 1.12 |
| 21. | Materials Science, MS | 37 | 1 | 34 | 33 | 11.78 | 10 | 6 | 8.51 | 0.72 | 1.30 |

* This data, according to Scopus, is for comparison. It confirms that Publish or Perish sees more (see Sec. 3).

Table 7
Matrix of significance evaluation of differences between the variances

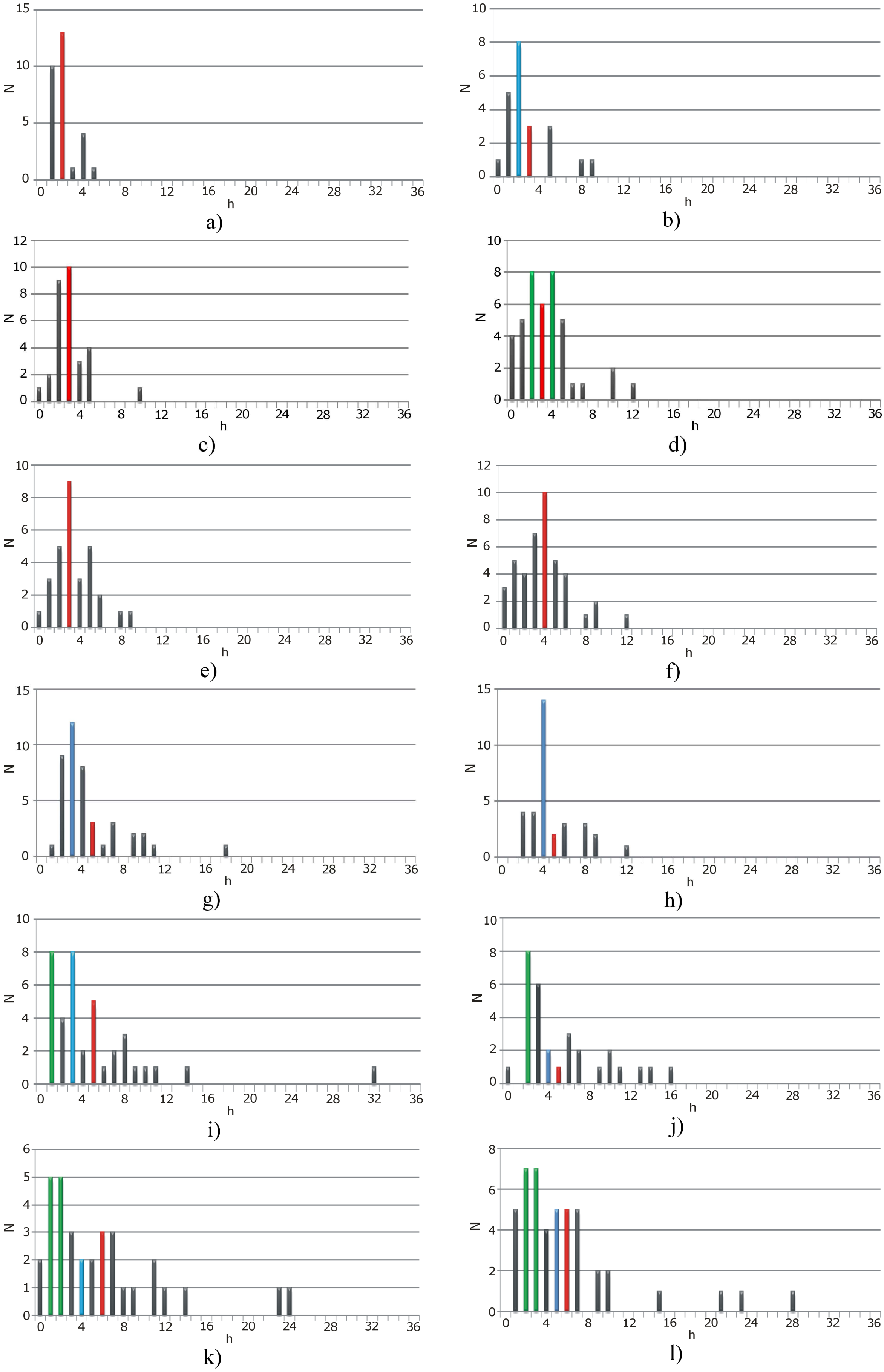
| | Discipline | A | IP | T | M | Sus | MB | CEN | G | Met | Env | Metr | Acou | TC | E | E&T | I | A&R | B&B | ChE | Mech | MS |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| others | A&U | | Y | | | | | | | | | | | | | | | | | | | |
| | IP | Y | | Y | Y | Y | | | | | Y | | | | | | | | | | | |
| lower medium cluster | Transport, T | | Y | | Y | Y | Y | | Y | | Y | | | | | | | | | | | |
| | Mining, M | | Y | Y | | Y | Y | Y | Y | Y | Y | | | | | | | | | | | |
| | Sustainable | | Y | Y | Y | | Y | Y | Y | | Y | | | | | | | | | | | |
| | MB | | | Y | Y | Y | | Y | Y | Y | Y | Y | | | | Y | | | | | | |
| | C&H | | | | Y | Y | Y | | Y | Y | Y | Y | Y | | | | | | | | | |
| | Geodesy, G | | | Y | Y | Y | Y | Y | | Y | Y | Y | Y | | | | | | | | | |
| upper medium cluster | Metallurgy, Met | | | | Y | | Y | Y | Y | | Y | Y | Y | Y | Y | Y | | | | | Y | |
| | Environmental | | Y | Y | Y | Y | Y | Y | Y | Y | | Y | Y | Y | Y | Y | Y | Y | | | Y | |
| | Metrology | | | | | | Y | Y | Y | Y | Y | | Y | Y | Y | Y | Y | Y | | | Y | |
| | Acoustic | | | | | | | Y | Y | Y | Y | Y | | Y | Y | Y | Y | Y | | | Y | |
| | T&C | | | | | | | | | Y | Y | Y | Y | | Y | Y | Y | Y | | | Y | |
| top cluster | EE | | | | | | | | | Y | Y | Y | Y | Y | | Y | Y | Y | Y | Y | Y | Y |
| | E&T | | | | | | Y | | | Y | Y | Y | Y | Y | Y | | Y | Y | Y | Y | Y | Y |
| | Informatics, I | | | | | | | | | | Y | Y | Y | Y | Y | Y | | Y | Y | Y | Y | Y |
| | A&R | | | | | | | | | | Y | Y | Y | Y | Y | Y | Y | | Y | Y | Y | Y |
| | B&B | | | | | | | | | | | | | | Y | Y | Y | Y | | Y | Y | Y |
| | ChE | | | | | | | | | | | | | | Y | Y | Y | Y | Y | | Y | Y |
| | Mech (*Scopus*) | | | | | | | | | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | | Y |
| | MS | | | | | | | | | | | | | | Y | Y | Y | Y | Y | Y | Y | |

Y – Yes, the hypothesis that the variances differ significantly may be rejected (on the basis of F-Snedecor test)

Table 8
The matrix of significance verification of differences between means

| | Discipline | A | IP | T | M | Sus | MB | CEN | G | Met | Env | Metr | Acou | TC | E | E&T | I | A&R | B&B | ChE | Mech | MS |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Others | A&U | | | | | | | | | | | | | | | | | | | | | |
| | IP | | | Y | Y | Y | | | | | | | | | | | | | | | | |
| Lower medium cluster | Transport, T | | Y | | Y | Y | Y | | Y | | | | | | | | | | | | | |
| | Mining, M | | Y | Y | | Y | Y | Y | Y | Y | | | | | | | | | | | | |
| | Sus | | Y | Y | Y | | Y | Y | Y | | | | | | | | | | | | | |
| | MB | | | Y | Y | Y | | Y | Y | Y | | Y | | | | Y | | | | | | |
| | C&H | | | | Y | Y | Y | | Y | Y | | | Y | | | | | | | | | |
| | Geodesy, G | | | Y | Y | Y | Y | Y | | Y | | | | | | | | | | | | |
| upper medium cluster | Metallurgy, Met | | | | Y | | Y | Y | Y | | Y | Y | Y | Y | Y | Y | | | | | Y | |
| | Env | | | | | | | | | Y | | Y | Y | Y | Y | Y | Y | | | | Y | |
| | Metr | | | | | | Y | | | Y | Y | | Y | Y | Y | Y | Y | Y | | | Y | |
| | Acou | | | | | | | Y | | Y | Y | Y | | Y | Y | Y | | | | | Y | |
| | TC | | | | | | | | | Y | Y | Y | Y | | Y | Y | | | | | Y | |
| top cluster | EE | | | | | | | | | Y | Y | Y | Y | Y | | Y | Y | Y | Y | Y | Y | Y |
| | E&T | | | | | | Y | | | Y | Y | Y | Y | Y | Y | | Y | Y | Y | Y | Y | Y |
| | Informatics, I | | | | | | | | | | Y | Y | | | Y | Y | | Y | Y | Y | Y | Y |
| | A&R | | | | | | | | | | | Y | | | Y | Y | Y | | Y | Y | Y | Y |
| | B&B | | | | | | | | | | | | | | Y | Y | Y | Y | | Y | Y | Y |
| | ChE | | | | | | | | | | | | | | Y | Y | Y | Y | Y | | Y | Y |
| | Mech (*Scopus*) | | | | | | | | | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | | Y |
| | MS | | | | | | | | | | | | | | Y | Y | Y | Y | Y | Y | Y | |

Y – Yes, the hypothesis that the mean values differ significantly may be rejected (on the basis of $t$-Student test)


N
h
a)
b)
c)
d)
e)
f)
g)
h)
i)
j)
k)
l)

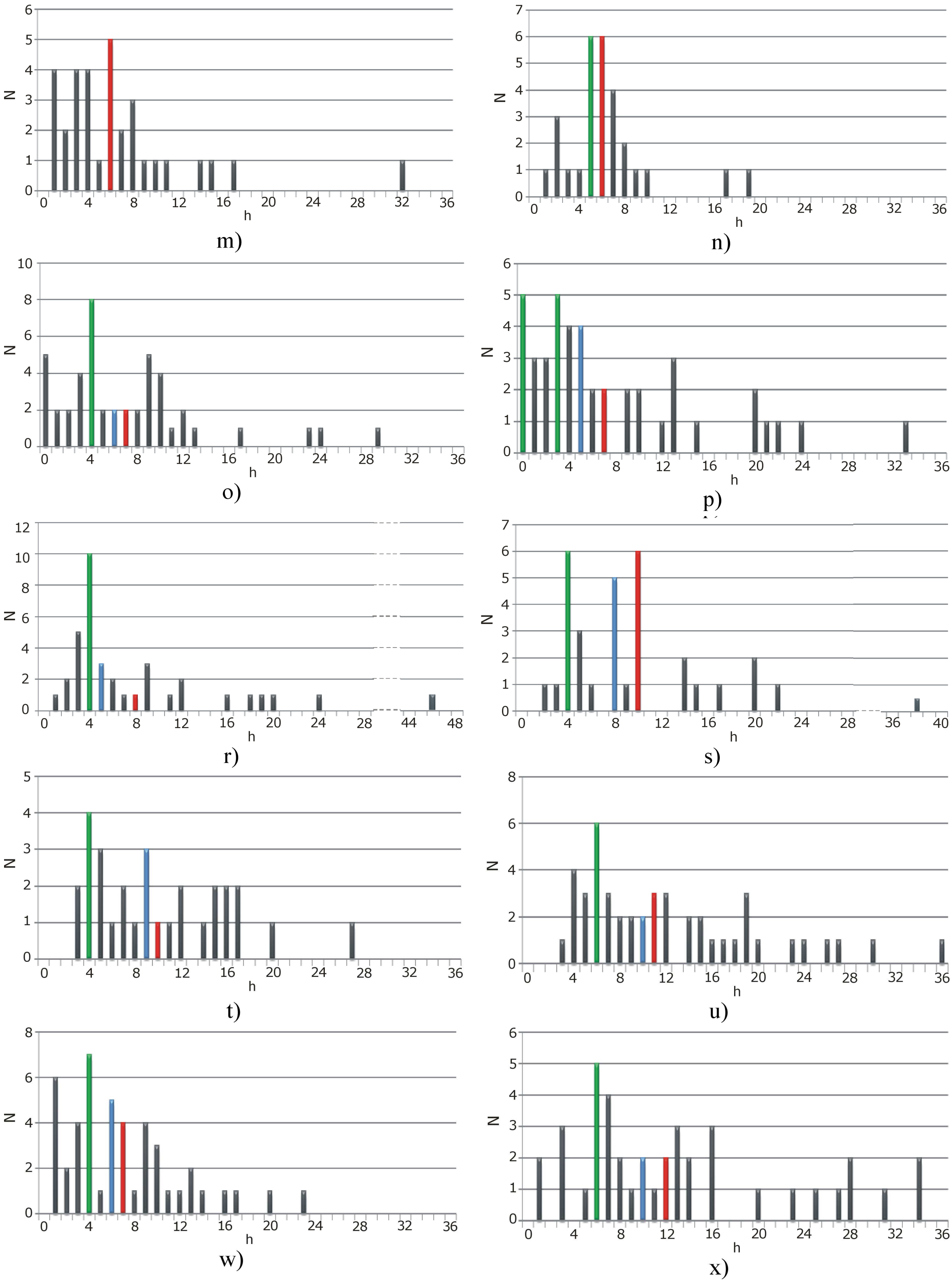



Fig. 12. Histogram of $h$-index distribution for various domestic representative populations, SRP: Architecture and Urban Planning (a), Production Engineering (b), Transport (c), Mining (d), Management of Mineral Resources (e), Machine Building (f), Civil Engineering and Hydroengineering (g), Geodesy (h), Environmental Engineering (i), Metallurgy (j), Metrology and Scientific Instrumentation (k), Electrical Engineering (l), Acoustics (m), Thermodynamics and Combustion (n), Electronics & Telecommunication (o), Informatics (p), Automatic Control & Robotics (r), Biocybernetics & Biomedical Engineering (s), Chemical and Process Engineering (t), Mechanics (u), Mechanics – Scopus (w), Materials Science (x); mean – red, mode – green, median – blue

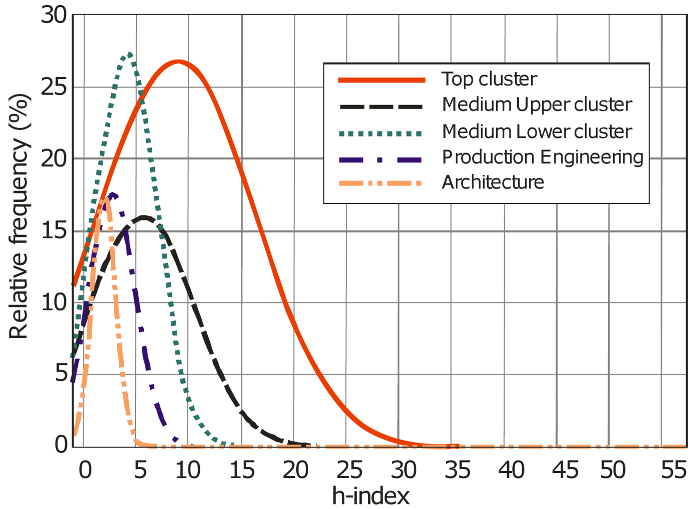



Fig. 13. Statistical distribution of $h$-index for top cluster, Medium Upper cluster, Medium Lower cluster, Production Engineering, and Architecture

Table 9
Restructured statistics measures of h-index distribution of domestic Engineering sub-disciplines

| No. | Engineering sub-discipline | Observation number, n | H | | | | | | Standard deviation $\sigma$ | Variation coeff. $\sigma/H$ | Skewnes coeff. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | min | max | range | mean | median | mode | | | |
| 1. | Top Cluster | 296 | 0 | 46 | 46 | 8.91 | 7 | 4 | 7.43 | 0.55 | 1.67 |
| 2. | Medium Upper Cluster | 178 | 0 | 32 | 32 | 5.64 | 5 | 2 | 4.97 | 0.88 | 2.27 |
| 3. | Medium Lower Cluster | 212 | 0 | 18 | 18 | 4.19 | 4 | 3 | 2.92 | 0.70 | 1.69 |
| 4. | Production Engineering | 22 | 0 | 9 | 9 | 2.82 | 2 | 2 | 2.28 | 0.81 | 1.54 |
| 5. | Architecture and Urban Planning, A | 29 | 1 | 5 | 4 | 2.07 | 2 | 2 | 1.13 | 0.55 | 1.13 |

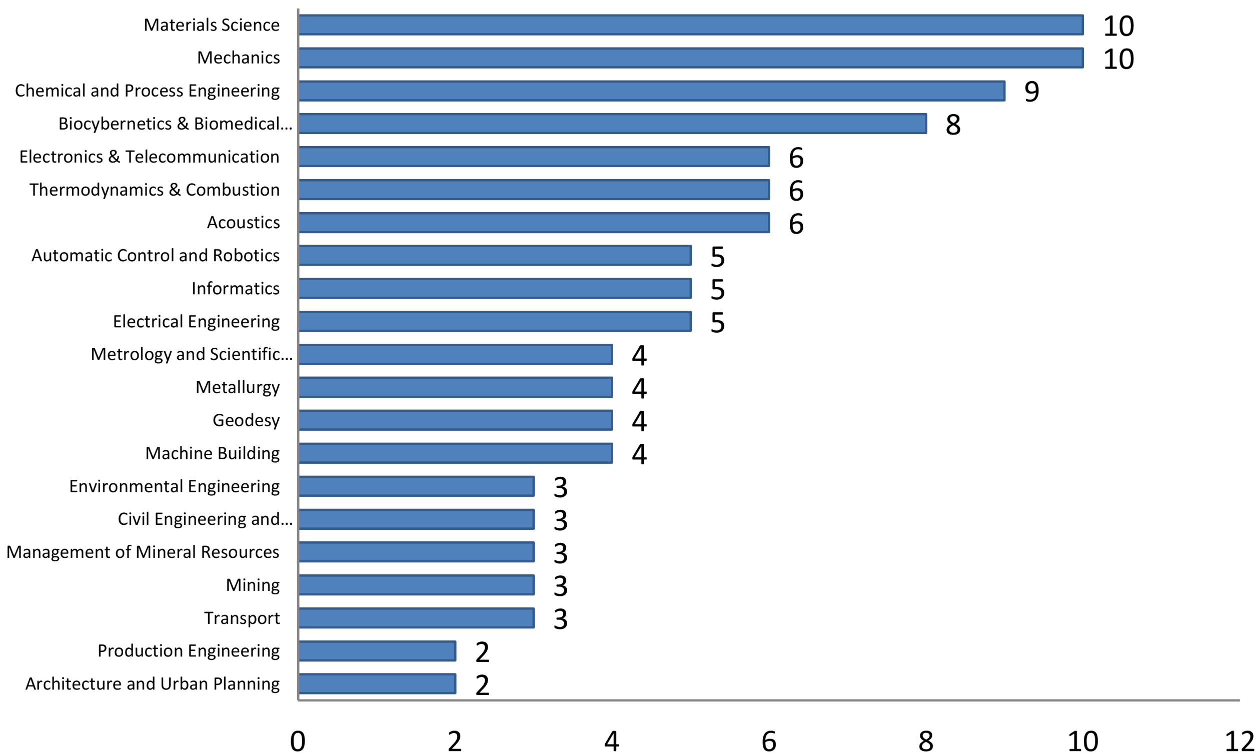



Fig. 14. The median – H-index value for domestic Engineering sub-disciplines

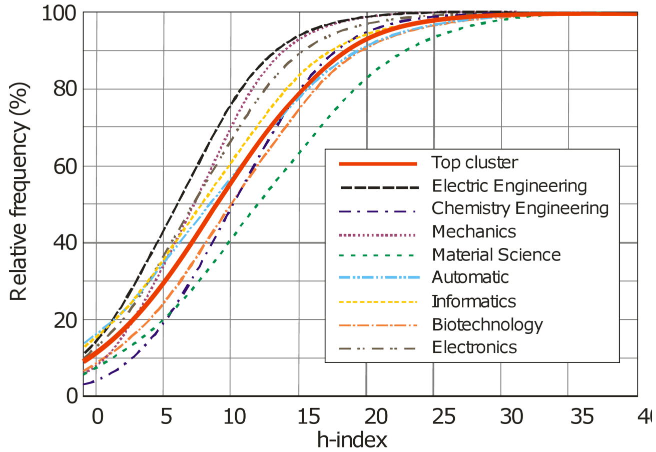



Fig. 15. Cumulative probability density function (CPDF) of $h$-index for Electrical Engineering, Chemical and Process Engineering, Mechanics, Material Science, Automatic, Informatics, Biotechnology, and Electronics. Joint CPDF for this "top cluster" is shown by red line

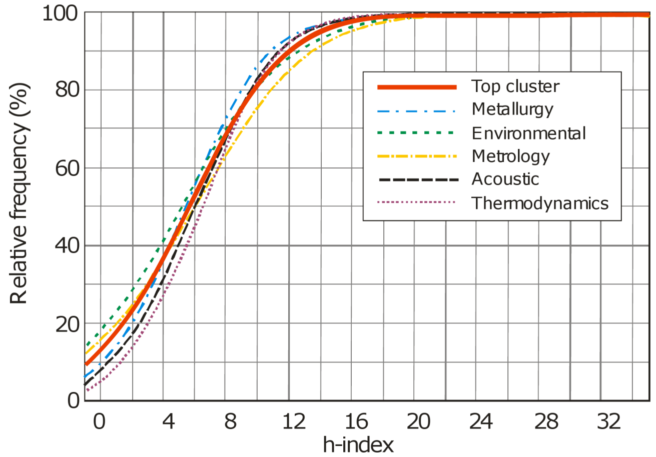



Fig. 16. Cumulative probability density function (CPDF) of $h$-index for Metallurgy, Environmental, Metrology, Acoustic, and Thermodynamics. Joint CPDF for this "upper medium cluster" is shown by red line

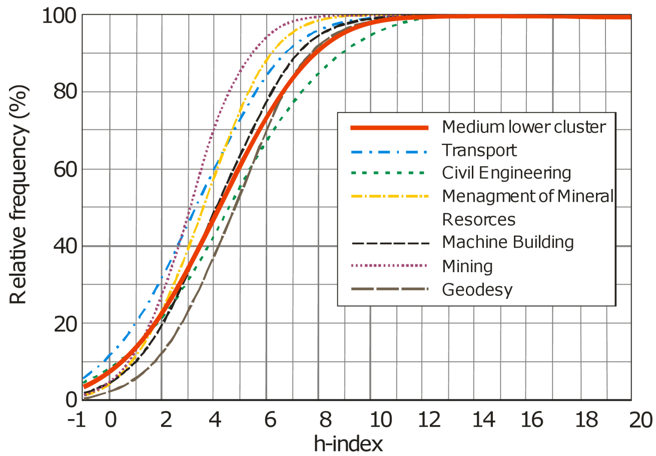



Fig. 17. Cumulative probability density function (CPDF) of $h$-index for Transport, Mining, Management of Mineral Resources, Machine Building, Civil Engineering, Geodesy. Joint CPDF for this "medium lower cluster" is shown by red line

As it has been already revealed the H-index of Engineering, as a whole, is undercounted and should be upgraded by the factor $f = 1.7$ (Sec. 4). However, internally among sub-disciplines of Engineering there also exist differences. It is particularly visible if we present the cumulative probability density function (CPDF) of the $h$-index for each engineering sub-discipline (Figs. 15–17). Someone could notice that in the Engineering only 30% of population show relative similarity in $h$-index. On the level above 30% of population, the difference between the top and medium clusters and others are significant.

It goes to the position that some sub-disciplines in Engineering need an additional normalization factor:

$$H_{e,s} = 1.7 \cdot f_s \cdot h, \tag{13}$$

where $f_s$ is the engineering sub-discipline factor. A defining $f_s$ value should be the subject of another study. It suggest that $f_s$ should be based on the characteristic median. There are medians (Fig. 18): 9:6:4:3 adequately for top, upper medium, lower medium and others. It seems that $f_s = 1.5$ and 2.3 for upper and lower medium cluster, respectively and 3 for other sub-disciplines have been already justified.

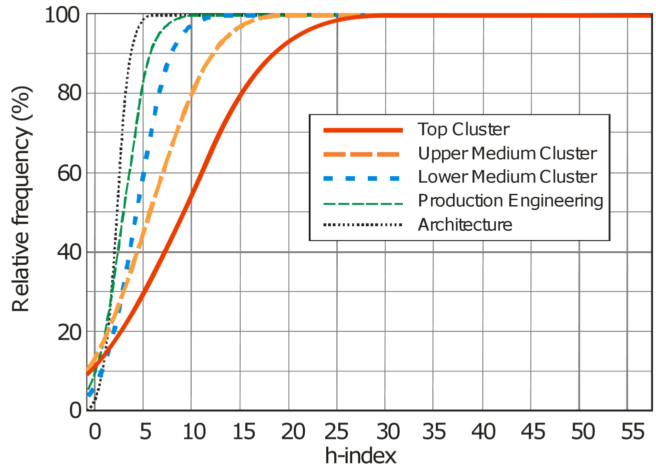



Fig. 18. Joint CPDF for clusters: top, upper medium, lower medium and "others": Production Engineering and Architecture

## 6. Conclusions

In general:

- It should be useful to recognize small letter $h$, as J.E. Hirsch originally suggested for individual researcher, and the capital H, as the collective indicator for journals, organizations, institutions and countries,
- published article should have enough time to be cited. The period of 7 years, since publication date, has been shown as needful for setting up stable $h$-value,
- $h$, H-indices are the ***general*** (commonly used) measure of the position of the given subject (person or organisation) and a change of the $h$,H-value, which never decreases in time, is a measure of progress. Guessing from the past the development of $h$,H-application seems to be inevitable. However, $h$,H-indices ***cannot be universal***. The challenge

for future is to find rationality between "necessity" and a potential limit which is comprised in the $h$,H-indices,

- there is a danger of confusion between the goal and a tool, or a tool and a game. If it is a game, the human nature should be a winner. The goal is the knowledge progress and $h$,H-indices are just the tool for measure, not otherwise,
- authors should have the right to free selection of a suitable database and machine for establishing the $h$-index. It is justified by the condition that unnoticed citation does not mean non-existing. In consequence, each $h$,H-value should be given with the adequate database/search machine,
- the position of Poland in the general country ranking and in the Engineering area is respectively 24 and 30. It is worse or even "much worse" in comparison with the Gross Domestic Product, GPD where Poland is on the $20^{th}$ position.

**In Engineering**:

- the citation ratio – the number of citations per paper could differ among various academic disciplines, even in the case of top scientists by the factor of 15, and between various Engineering sub-disciplines 10 times,
- the H-index of Engineering, as a whole, is undercounted, when compared to other academic disciplines. Even the top scientists in Engineering should publish 2 times more papers to receive the same $h$-index as top scientists in overall area. There are various nominalization factors for Engineering relative to overall disciplines or to Physics:

| Reference | $h = f_h \cdot h_E$ | $N = f_N \cdot N_E$ | |
|---|---|---|---|
| | $f_h$ | $f_N$ | |
| W. Glänzel [34, 35], J. Podlubny [31] | 2.5 | 2.0 | to Physics |
| J.E. Iglesias and C. Pecharroman [2] | 1.7 | 2.5 | to overall disciplines |
| in this paper is according to "the best" – section 6 | 1.6 | 2.4 | to overall disciplines |

The compensation factor equals to 1.7, suggested by Iglesias and Pecharroman [2] seems to be utterly justified,

- the Polish Engineering sub-disciplines, similar to world situation, show differentiation in H-index in the extreme case by the factor of 5. The Engineering sub-disciplines are categorized into four groups and the compensation factors are proposed, when similarity with the Engineering situation versus other academic disciplines is considered.
- Concluding, on the base of H-median, the Polish Engineering sub-disciplines represent at least 30% of an adequate world value. However, in many Engineering sub-disciplines there exist outliers with $h$-index 2 times or more higher than the world median. There are many reasons for undercounting the Polish Engineering: country-of-origin effect, non-English authors, publishing (partially) in domestic journals written in the Polish language could be mentioned among others.

**Acknowledgements.** Authors would like to express acknowledgements to Joanna J. Sokołowska, Ph.D. and Piotr Woyciechowski, Ph.D. from Warsaw University of Technology for gathering basic data.

REFERENCES

[1] L. Czarnecki, "Challenges towards civil engineering environment in a light of scientific policy assumptions" *Engineering and Construction* 9, 499–502 (2011), (in Polish).
[2] J.E. Iglesias and C. Pecharroman, "Scaling the h-indexes for Different Scientific ISI Fields", *Scientometrics* 73 (1), 303–320 (2007).
[3] www.beggerbrains.com
[4] E. Ecco, *The Infinity of Lists*, Mac Lehose Press, London, 2009
[5] B.E. Lautrup "Measure for measures", *Nature* 444, CD-ROM (2006).
[6] M. Porter, "Trust in numbers: the pursuit of objectivity in science and public life", Princeton University Press, Princeton, 1995.
[7] C.K. Życzkowski ,"Citation graph, weighted impact factors and performance indices", *Scientometrics* 85, 301–315 (2010).
[8] A. Sawicki ,"Picture of the Polish science after MNiSW ranking", *Academic Forum* 11, CD-ROM (2010), (in Polish).
[9] P. Heymann, A. Paepcke, and H. Garcia-Molina, "Tagging Human Knowledge", http://ilpubs.stanford.edu:8090/956/2/ltoc-super-preprint.pdf (2009).
[10] A.W. Harzing and R. van der Wall ,"Comparing the Google Scholar h-index with the ISI Journal Impact Factor", *Research in Int. Management Products and Services for Academics* 1, www.harzing.com/h_indexjournals.htm (2008).
[11] A.W. Harzing and R. van der Wall, "A google scholar h-index for journals: a better metric to measure journal impact in economics and business", www.harzing.com (2007).
[12] J.K. Vanclay, "Impact factor: outdated artifact or stepping-stone to journal certification", *Scientometrics* doi:10.1007/s11192-011-0561-0 (2012).
[13] R. Rousseau,"Updating the journal impact factor or total overhaul?", *Scientometrics* doi: 10.1007/s11192-012-0649-1 (2012).
[14] E. Garfield, "Citation indexes to science: a new dimension in documentation through association of ideas", *Science* 122 (3159), 108–111 (1955).
[15] E. Garfield, "The agony and the ecstasy – the history and meaning of the journal impact factor" *Int. Congress on Peer Review and Biomedical Publication* 1,www.eugengarfield.org (2005).
[16] A. Rogalski, M. Kaźmierkowski, and L. Czarnecki, "Technical sciences in international rankings", *Academic Forum* 9, 34–38 (2012), (in Polish).
[17] A. Osielawska,"Evaluation measures of journals and scientists", *EBIB* 8, CD-ROM (2008), www.ebib.info/2008/-99a.php?osielawska
[18] Kuan-Teh Jeang "Impact factor, H-index, peer comparisons and retrovirology: is it time to individualize citation metrics?", *Retrovirology* 4 (42), www.retrovirology.com/content/4/1/42 (2007).
[19] N. Adler and A.W. Harzing,"When knowledge wins: transcending the sense and nonsense of academic ranking", *Academy of Management Learning and Education* 1, www.harzing.com (2009).

[20] P.A. Lawrence "Lost in publication: how measurement harms science" *Ethics Sci. Environ. Polit.* 8, 9–11 (2008).

[21] P.A. Lawrence "Rank injustice: the misallocation of credit is endemic in science", *Nature* 415, 835–836 (2002).

[22] J.E. Hirsch "An index to quantify an individual's scientific research output", *Proc. Nat. Acad. Sci.*, *PNAS* 102, 46: 16569 (2005).

[23] R. Kierzek "Polish science in Hirsch index", *Science Matters* 6–7, 137 (2008), (in Polish).

[24] J.E. Hirsch "Does the h-index have predictive power?", *Proc. Nat. Acad. Sci.* 2, 7080646 (2007).

[25] *SJR – Scimago J. and Country Rank Retrieved*, http://www.scimajor.com (2012).

[26] G.J. Woeginger, "An axiomatic characterization of the Hirsch index", *Material Social Sciences* 56, 224–232 (2008).

[27] M. Kosmulski, "I-a bibliometric Index", *Academic Forum* 11, 31 (2011), (in Polish).

[28] G. Prathap, "Hirsch – type indices for ranking institutions scientific research output", *Current Science* 91 (11), 1439 (2006).

[29] P.D. Batista, "Is it possible to compare researcher with different scientific interests?", *Scientometrics* 68 (1), 179–189 (2006).

[30] E. Lillquist and S. Green, "The discipline dependence of citation statistics", *Scientometrics* 84, 749–768 (2010).

[31] I. Podlubny, "Comparison of scientific impact expressed by the number of citations in different fields of science", *Scientometrics* 64 (1), 95–99 (2005).

[32] K. Sangwal "On the relation between citation of publication output and hirsch index h of authors: conceptualization of tapered hirsch index $h_t$, circulation area radius r and citation acceleration", *Scientometrics* 93, 987–1004 (2012).

[33] S. Alonso, "H-index: a review focused on its variants, computation and standardization for different scientific fields" *J. Informetrics* 3, 273–289 (2009).

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

[35] W. Glänzel "H-index concatenation", *Scientometrics* 77 (2), 369–372 (2008).

[36] E. Csajbok, "Hirsch index for countries based on essential science indicators data", *Scientometrics* 73 (1), 91–117 (2007).

[37] *Microsoft Academic Search,* http://academic.research.microsoft.com/.

[38] *Chemistry World* www.rsc.org/chemistryworld/News/2007/-April/23040701.asp.(2011).

[39] *The h-index for Computer Science* www.cs.ucla.edu/-palsberg/h-numi/.

[40] G. Gugliota "The genius index: one scientists crusade to rewrite reputation rules", *Wired Magazine*, www.wired.com/culture/geekipedia/magazine/17-06/mf-impactfactor?-currentP/(2009).

[41] A. Sidiropoulous, D. Katsaros, and Y. Monolopoulos, "Generalized h-index for revealing latent facts in social networks of citations", *Scientometrics* 72 (2), 253–280 (2007).

[42] M. Żylicz, "Parametric evaluation today and tomorrow", *Academic Forum* 6, http://www.forumakad.pl/archiwum/2006/06/22_ocena_parametryczna.html (2006).