Evaluation of publications by the part-set method

Péter Vinkler[1]

Abstract

Several scientists publish journal papers on more than only a single scientific field or subfield. However, in evaluating the impact of publications of individuals their all scientific publications are taken into account, in general. In this paper a new method, namely the *part-set method* is introduced for assessing papers according to the respective fields and forming a *cumulative impact* index. In assessing publications of scientists, *elite set* indicators (e.g. *h*-index, *g*-index, top percentage indices, π-index) are preferably used recently. For obtaining elite sets, we may apply *inside* standards, i.e. values derived from the number of papers and citations in the set analysed, or *field* (outside) standards taking into account the same factors referring to all papers on the corresponding field. For model calculations some scientometricians are selected who publish papers not only on scientometrics but also on other (e.g. physics, chemistry, medicine, etc.) fields. Consequently, their papers form *complex* sets of which *part*-sets may show different bibliometric features. In this publication, the number of papers in the elite set of the scientometric part-set and their citations are presented using both inside and outside standards. No significant correlation was found between the number of *papers* in the elite sets obtained by inside and that calculated by field standards. The number of *citations* in the different elite sets calculated by both inside and outside standards however, significantly correlates with each other. The presented model indicates that the sum of *h*-index of two part-sets derived from a common complex (total) set can be *equal to* or *higher* than the *h*-index of the corresponding complex set. In contrast, the sum of π-index of two part-sets can be *higher* or *lower than* or *equal to* the π-index of the parent complex set. The model reveals that the *maximum* value of the sum of *h*-index of two part-sets belonging to the same complex set can be two times the *h*-index of the corresponding complex set. In evaluating *total* scientific impact (e.g. life-work) of publications of scientists who are active in several fields, the application of *sum* (or weighted sum) of the pertinent impact indices (e.g. *h* or π-index) obtained for the individual fields *separately* may be recommended instead of calculating the impact of the total set.

✉ Péter Vinkler
vinkler.peter@ttk.mta.hu

1   Research Centre for Natural Sciences, Hungarian Academy of Sciences, Magyar Tudósok krt. 2,
    1117 Budapest, Hungary

Introduction

After introducing the *h*-index (Hirsch 2005), the methods calculating impact indices from an appropriately selected *smaller set* out of the total, are preferably applied in assessing publication impact of scientists or journals (Vinkler 2009, Schreiber 2010, Bornmann et al. 2011, Schreiber et al. 2012, Wildgaard et al. 2014, Todeschini and Baccini 2016). The appropriately selected smaller sets may be termed as *elite or core* sets (Vinkler 2017a, b, 2019).

The application of elite sets involves the assumption that indicators derived from an *exclusive part* of the publications with highest number of citations may represent the relative impact of the results more appropriately than that obtained from the total. The mentioned practice is in accordance with one of the basic paradigms of scientometrics: *Citations indicate impact*, and *more citations indicate greater impact*.

It is well known that the distribution of citations over journal papers is skewed (Seglen 1992). Accordingly, the highly cited papers in the elite set may be regarded as relatively most influential within the corresponding total set. There are several methods for obtaining elite sets, e.g., *h*, *g*, and *π*-statistics, or applying a certain *percentage* of the total (Vinkler 2007, 2019).

J. Hirsch, who introduced the *h*-index (Hirsch 2005), is a university professor of physics. He published altogether 272 journal papers up to 03 06 2019. Until that date, his papers received 19,304 citations. The mean citation frequency of the papers is equal to 71.23. Taking into account all journal papers of Hirsch, independent of their topic, his *h*-index is 58 (see WoS). The mentioned author initiated a revolution in the development of scientometric impact indices by suggesting the *h*-index.

Hirsch's first scientometric paper (Hirsch 2005) obtained 3633 citations up to the mentioned date. Accordingly, we would think, he was one of the most influential scientists in the field. However, surveying the publication list of the mentioned author in WoS, among his 272 papers only 5 articles may be classified as *scientometric* publications. The citation rate of these papers is as follows: *1*/(published in 2005): 3633, *8*/(2007): 434, *31*/(2010): 123, *145*/(2014): 25, and *249*/(2019): 0. (The cursive numbers are the rank numbers of the papers by citation taking into account all papers (272) of the mentioned author.) Accordingly, the *h*-index, i.e. the *scientometric impact* of J. Hirsch would be equal to 4 which seems to be a rather low value. It is because the value of *h*-index *cannot exceed* the number of publications in the set studied, and it does not regard the total number of citations obtained (Vinkler 2007, 2017a).

In contrast, calculating the *π*-index (Vinkler 2009) of the *scientometric* papers of Hirsch, a relatively high value is obtained. Hirsch published five scientometric papers up to 03 06 2019. Accordingly, the number of his *π*-core papers: $\sqrt{5} = 2.24$ (rounded as 2), see Table 1. In calculating the *π*-index, we have to sum up the number of citations to his *π*-core papers, and the sum should be divided with 100. Accordingly Hirsch's *π*-index = 0.01 $(3633 + 434) = 40.67$. This value seems to be rather high compared to that for well-known scientometricians (see Table 1 and Vinkler 2017b). (In Table 1 the *π*-index of papers of Hirsch in the scientometric part-set is 39.81. The reason for the difference (40.67 vs. 39.81) is the different date (03 06 2019 vs. 30 01 2019) for obtaining the citation data.)

The given example makes it clear that the *π*-index prefers scientists with high number of citations to *π-core* publications even they have published relatively low number of papers.

The above example makes it reasonable to *separate* publications in the complex set of researchers working on more than a single science field by part-sets when evaluating their

Table 1 Number of publications (P), h-index (h), and π-index (π) for the total set (T) and its elite part-sets (Scientometrics: S and other sciences: O, respectively) of some scientometricians

Name	Total set			Part sets		Part sets		Part sets		Sum of part sets	
	P (T)	h (T)	π (T)	P (S)	P (O)	h (S)	h (O)	π (S)	π (O)	H (S+O)	Π (S+O)
Abt HA	189	30	24.18	61	128	13	29	2.10	22.66	42	24.76
Braun T	263	32	19.79	150	113	26	22	15.85	6.71	48	22.52
Hirsch JE	272	58	105.62	5	267	4	57	39.81	110.21	61	150.02
Kosmulski M	177	32	23.63	14	163	5	32	0.56	23.63	37	24.19
MacRoberts MH	30	12	9.33	21	9	11	5	9.33	0.53	16	9.86
Moravcsik MJ	165	16	8.21	88	77	11	13	6.07	2.98	24	9.05
Pouris A	98	15	3.60	69	29	14	6	2.54	1.49	20	4.03
Seglen PO	181	61	103.54	11	170	9	58	16.99	92.39	67	109.38
VanRaan AFJ	54	23	10.87	49	5	23	4	10.87	0.30	27	11.17
Vinkler P	74	24	6.39	59	15	22	8	4.91	2.48	30	7.39
Mean	150.3	30.3	31.52	52.7	97.6	13.8	23.4	10.90	26.34	37.2	37.23
SD	83.8	16.9	39.16	44.2	86.8	7.5	20.5	11.61	40.66	17.2	50.20

$H = h(S) + h(O)$
$\Pi = \pi(S) + \pi(O)$

activity. Separating publications according to research fields finds support in the fact that the *bibliometric features* (i.e. mean number and age of references in papers, yearly number of the publications, etc.) are rather different by the field (Vinkler 2010). It is obvious that criteria for selecting the papers in a complex ("mixed") set by parts should be determined according to the goals of the assessment.

As criteria for classifying the publications within a complex set, different publishing *periods*, *cooperating* partners or *affiliations* of the authors, etc. may also be applied.

As *complex publication sets*, the sets consisting of two or more part-sets of publications with different bibliometric features may be regarded. From this definition of complex sets it follows that a publication *part-set*, i.e. part of a complex publication set selected by scientific disciplines or fields or subfields (or other criteria) would contain bibliometrically (or by other aspects) more or less consistent publications.

It seems to be obvious, if we were interested in determining the impact of publications of a physicist on physics, we should not calculate with his or her possible other publications on e.g. economics or history. In addition, for comparing scientific production of individuals working on different fields, we have to select different reference standards. Anyway, the *goals* of the evaluation should determine the possible classification of the papers in complex sets.

The methods, how to obtain *elite sets* may be distinguished according to the standards applied. The calculations may apply *inside* or *outside* standards.

Inside publication standards are reference values for publication assessments calculated from the data of the studied set. Such value is, e.g. the *h*-index of the studied publication set. This indicator represents the *number of papers* in the *h*-core. We may use also the number of π-core papers as elite papers, which is equal to the square root of total papers, or we may calculate the number of those papers in the studied set, that obtained more than an arbitrary number (e.g. 50 or 100) of citations. There are of course also other possibilities. E.g. the number of papers that are cited more heavily than the mean or the most cited 10% of papers in the set, etc.

Outside (field) standards are reference values obtained from data of a set selected as reference. Such reference set may be e.g., the set of papers in a single or in several journals, or in all journals in the field of which topic is related to that of the papers studied. As standard e.g., the set of papers with higher number of citations than the mean citation frequency (C/P) of the papers in a selected journal or in journals of the field may be applied. Accordingly, we may calculate *the number of publications from the set studied within the reference set* of papers selected according to the mentioned method.

The aims of the present study:

- comparing the number of journal *papers* in the elite publication sets selected by different methods and comparing the number of *citations* to those papers,
- comparing the *h* and π index calculated for complex (total) publication sets with the *sum of h* and π indices, respectively obtained for two part-sets within the corresponding complex set, separately.

Methods and data

Model 1 in Table 2 demonstrates one of the main research questions of the present study. It is assumed that the compl**E**x publications set (E) of a scientist consists of 8 publications, which can be attributed to two different subject fields: a1 and a2. The papers are listed by

decreasing number of citations. The *h*-index of set E is four because there are four papers with four or more citations. Let us suppose, the part-sets (a1 and a2) derived from the parent E set, would contain five and three papers, respectively. Accordingly the *h*-index for set a1 = 3 and for set a2 = 2. The sum of *h*-index of the part-sets: $h(a1) + h(a2) = 5$ that is higher than the *h*-index for the parent set: $h(E) = 4$. This observation would indicate to distribute set E into part-sets, if we were interested in impact of the papers on field a1 or on field a2, separately.

For the present study, 10 scientists were selected who published papers both in scientometrics (S) and in other field(s) (O) (Table 1). Scientometrics was selected because the author is active on this field for several years. The publications were classified as *sciento-metric* and "other" according to their title and abstract. Each abstract was surveyed individually if the title of the paper did not yield enough information for the decision. Consequently, not only publications in journals classified as library and information science (LIS) by WoS were taken into account as scientometric publications, but also those papers in any journal (e.g. Research Policy, Nature, Analytical Chemistry, etc.) which could be classified as scientometric. In addition, of course, not all papers published in LIS journals were accepted as scientometric publications. The papers were collected by the name of the authors from WoS. Both their scientometric, bibliometric, webometric and altmetric papers were classified as "scientometric". Papers e.g. on physical, chemical, astronomical, astrophysical, medical, and biological topics and on library science were attributed to "other" fields.

The number of and citations to the papers were collected from Clarivate Analytics Web of Science All Databases 1975–2018 on 30 January, 2019. The *field standard* of citation frequency (Cf/Pf) was calculated from data of Scientometrics on 01 September, 2019. In 1975–2018 Pf = 5,542 papers were published in the periodical to which 87,409 citations (Cf) were received in the same period. Accordingly the citation frequency: Cf/Pf = 15.77 (rounded: 16). The standards applied in the paper, i.e. number of citations by paper are the following: 2.5Cf/Pf = 40 and 5Cf/Pf = 80. The paper in Scientometrics ranked as (P/100)th = 55, obtained

Table 2 Model 1: Calculating the sum (H) of *h*1-index and *h*2-index of two part-sets (a1 and a2, respectively) from the same parent set (E)

Rank of publications (*i*)	E $c(i)$	a1 $c(i)$	a2 $c(i)$
1	8	8	
2	7		7
3	6	6	
4	5		5
5	4	4	
6	3	3	
7	2	2	
8	1		1
Index			
h(E)	4		
h(a1)		3	
h(a2)			2

c_i: number of citations to the *i*th publication

$h(E) = 4$

$H = h(a1) + h(a2) = 3 + 2 = 5$

150 citations in the mentioned period. Accordingly, those papers of the studied scientists that obtained more citations than 150 are counted as papers belonging to the elite set of this type.

For obtaining *inside* standards (i.e. standards calculated from data of the publication set analysed) the following methods were applied:

Methods taking into account *the size of the publication set* analysed (P: number of publications) for obtaining the number of papers in the elite set (Pe):

- *π-statistics*: $P(e) = P(\pi) = \sqrt{P}$, or $2P(\pi) = 2\sqrt{P}$, or $\sqrt[3]{P}$
- *logarithmic statistics:* $P(e) = 2 \log P$
- *percentage statistics:* $P(e) = 0.1P$

A method taking into account the *rank of papers* by citation and *number of citations* obtained to the individual papers:

- *h*-statistics: $P(h)$, i.e. number of papers in the *h*-core.

Methods applying *outside* standards (here: field standards, e.g. those papers published in a special journal or all papers published on the corresponding field during a selected period):

- Number of papers in the set studied with number of citations 5 times (or 2.5 times) higher than the *mean citation rate* (Cf/Pf = 80 and 40, respectively) *of journal papers in Scientometrics* in the studied period. (Cf is the total number of citations whereas Pf is total number of papers.)
- Number of papers in the analysed set with equal to or higher than 150 citations.

It is obvious that studying relations between part sets and complex sets several indicators may be selected. In this study two elite set indicators, namely *h*-index and *π*-index are chosen. The former may represent most *h*-type indices, whereas *π*-index may refer to the elite set indices, which are obtained by simple mathematical calculations (e.g. 1% or 10% of total papers). Table 3 makes it possible to compare the number of papers in *h*-core and *π*-core with the size of elite sets obtained by other methods.

The author was not interested in attributing papers to other topics or disciplines outside scientometrics. This way the present study may refer to „two part-sets systems", i.e. papers *within* a complex (mixed) set consisting of publications on scientometrics and on „other" topics. The study with more than two part-sets may be more complicated. The „two part-sets system" presented here may be regarded as the starting step toward evaluating more complex publication systems of scientists through separating the publications according to different topics, time-periods, affiliations, cooperating partners, etc.

The statistical analyses were performed with STATISTICA data analysis software system version 13, TIBCO Software Inc. Palo Alto, USA.

Results

Complex sets and part-sets

Table 1 shows the number of publications (P), *h*, and *π*-index of the scientometricians analysed. The number of papers in the *complex* (total) set ranges from 30 (MacRoberts)

Table 3 Number of *papers* in different elite sets (P(h), P(π), etc.) of the scientometric publications (part-sets S in Table 1) of some scientometricians

Name	Number of papers								
	Applied *inside* standards						Applied *field* standards		
	P(h)	P(π)	2P(π)	2logP	$^3\sqrt{P}$	0.1P	0.01Pf $c_i \geq 150$	5Cf/Pf $c_i \geq 80$	2.5Cf/Pf $c_i \geq 40$
Abt HA	13	8	16	4	4	6	0	0	0
Braun T	26	12	24	4	5	15	3	8	8
Hirsch JE	4	2	4	1	2	1	2	3	3
Kosmulski M	5	4	8	2	2	1	0	0	0
MacRoberts MH	11	5	10	3	3	2	2	4	7
Moravcsik MJ	11	9	18	4	4	9	1	1	2
Pouris A	14	8	16	4	4	7	0	0	3
Seglen PO	9	3	6	2	2	1	3	5	7
VanRaan AFJ	23	7	14	3	4	5	3	8	18
Vinkler P	22	8	16	4	4	6	0	1	7
Mean	13.8	6.6	13.2	3.1	3.4	5.3	1.4	3.0	5.5
SD	7.5	3.1	6.1	1.1	1.1	4.4	1.3	3.2	5.3

P: total number of scientometric (S) papers (see Table 1)

Pf: total number of papers in the field

Cf: number of citations to Pf papers

$P(\pi) = \sqrt{P}$

up to 263 (Braun) and 272 (Hirsch). The ratios of the number of papers in the *part* sets (scientometrics, S related to "other" fields, O) range from 0.02 (Hirsch) and 0.06 (Seglen) up to 3.93 (Vinkler) and 9.80 (VanRaan). There are individuals publishing papers almost exclusively on non-scientometric topics (e.g. Hirsch, Seglen, Kosmulski), whereas other scientists publish similar number of papers in both scientometrics and in other fields (e.g. Braun or Moravcsik). There are researchers with many more scientometric papers than publications on other topics (VanRaan and Vinkler).

It is worth mentioning that the Spearman correlation coefficient between the number of papers (P) in the complex set and *h*-index of the studied scientists is 0.79, and that between P and π-index: 0.72, whereas the correlation coefficient between *h*-index and π-index is 0.83. The mentioned coefficients are significant at $p \leq 0.05$.

Both *h* and π-index reflect scientific excellence. Both indices depend on the number of publications and citations and on the *distribution* of citations over publications. Nevertheless, the *h*-index lays greater stress upon the number of publications than the π-index. According to Iglesias and Pecharromán (2007) and Schubert and Glänzel (2007) there is a strong linear correlation between *h*-index and the product: [$k \cdot$(C/P)$^{2/3} \cdot$P$^{1/3}$], where C is the total number of citations to P papers in the set analysed and *k* is a constant. In contrast, the π-index lays greater stress upon the number of citations to the *most frequently* cited papers than the *h*-index, as it calculates with the number of citations to the top P$^{1/2}$ papers multiplied by 0.01.

Table 1 shows both the *h*-index and π-index of papers of the studied scientists also by *part-sets*. There are scientists with significantly higher indices for their

non-scientometric publications, h(O) or π(O) > h(S) or π(S), respectively, like Abt, Hirsch, Kosmulski, and Seglen, whereas other individuals (Braun, MacRoberts, Pouris, VanRaan, and Vinkler) show higher indices for their scientometric papers than that for other topics. Moravcsik seems to be an exception. His h-index is higher for papers in non-scientometric field (theoretical physics) than that in scientometrics (O: 13 vs. S:11). In contrast to this, his π-index is lower in physics than in scientometrics (O: 2.98 vs. S: 6.07). Taking into account the similar number of papers in the part-sets (S: 88 vs. O: 77), the higher π-index may indicate a greater influence on scientometrics than on theoretical physics. Naturally, for drawing the mentioned conclusion, we should assume similar bibliometric features for both fields.

The difference between the h(O) and h(S) or π(O) or π(S) should be influenced by the different number of papers in the part-sets. The scientists publishing more scientometric papers than papers on other topics (Braun, MacRoberts, Pouris, VanRaan, and Vinkler) show higher, whereas that with less papers in scientometrics than on other fields (Abt, Hirsch, Kosmulski, and Seglen) lower h(S) and π(S)-index than h(O) and π(O), respectively.

Table 1 contains also means and standard deviation (SD) values of the indicators. Surveying the SD values it is obvious that the differences between means cannot be significant in most cases. One of the reason that the number of cases (10) is relatively low. The difference is significant however between the mean number of papers in the complex set, P(E) and mean h(E) ($p = 0.0003$) and mean π(E) ($p = 0.0007$). Further, the difference is significant between the mean h-index of the total set, h(E) and mean h-index of the scientometric part-set, h(S) ($p = 0.011$).

No significant difference was found between the *mean h and π-index of the complex set*, h(E) and π(E) and the corresponding mean of *sum of indices* of the part-sets, H(S + O), Π(S + O), respectively. Nevertheless it is not surprising taking into consideration the high SD values (Table 1). The high SD values may be caused by the following factors: different bibliometric features of the fields where the authors are active, different number of papers of the individuals on the different fields, and the relatively low number of scientists studied.

In contrast to the mentioned general tendency of significance, it seems to be relevant to take into account the difference between the h or π indicator referring to the complex (total) set and the sum of the indicators derived as cumulative number of the index of the corresponding part-sets on *individual level*. Naturally, the difference mentioned may refer to any other impact indicator.

For clearing the relations between the mentioned indices, Model 2 and 3 would yield some more information.

Calculating elite subsets by different methods

The size of the elite sets of scientometric papers of the scientists obtained by inside and outside standards are shown in Table 3. The data indicate the following conclusions:

- P(h) > P(π)
- 2P(π) > P(π)
- P(h) and P(π) > 2logP

Earlier studies maintain the statement: P(h) > P(π) (Vinkler 2011, 2017a, b). Relations: 2P(π) > P(π) and P(π) > 2logP are obvious.

Figure 1 shows the *percentage rate* of papers in the elite sets related to the total number of scientometric papers of the studied scientists. The rank of scientometricians on the x-axis reflects the decreasing number of citations in the h-core of their papers.

The highest share can be obtained in most cases by *h*-statistics. The dynamic range of P(*h*) indices: 12.50% (Moravcsik) – 81.82% (Seglen). The shares obtained by 2P(π) method show similar values: 16.00% (Braun) - 80.00% (Hirsch). The percentage rate strongly depends on the distribution of citations over the papers.

The Spearman correlational coefficients of the indices are given in Table 4. The data show, the number of scientometric papers, P(S) in the set *correlates* significantly (at $p \leq 0.05$) with the size of elite sets obtained *by inside standards*: P(*h*), P(π), 2P(π), 2logP, $^3\sqrt{P}$, 0.1P ($r = 0.73$–0.99). However, P(S) *does not correlate* significantly with the size of elite sets obtained *by outside standards* ($c_i \geq 150$, $c_i \geq 80$, and $c_i \geq 40$). The size of the elite sets obtained by inside standards correlates significantly with each other (0.73–1.00). The same is valid also for the correlation between 5Cf/Pf and 0.1Pf ($r = 0.96$), 2.5Cf/Pf and 0.1Pf ($r = 0.74$), and 2.5Cf/Pf with 5Cf/Pf ($r = 0.88$). All the other correlations between the sizes of elite sets calculated by with outside standards are not significant at $p \leq 0.05$.

The size of elite sets (i.e. number of papers in the elite sets) should not be regarded as an impact measure, a priori. The number of papers in the π-core, i.e. P(π) of the publication set of individuals shows only *quantitative* aspects. However, the *size* of the *h*-core, P(*h*) (i.e. number of papers in the *h*-core) *is equal to the h-index* which may be regarded as an *impact* index.
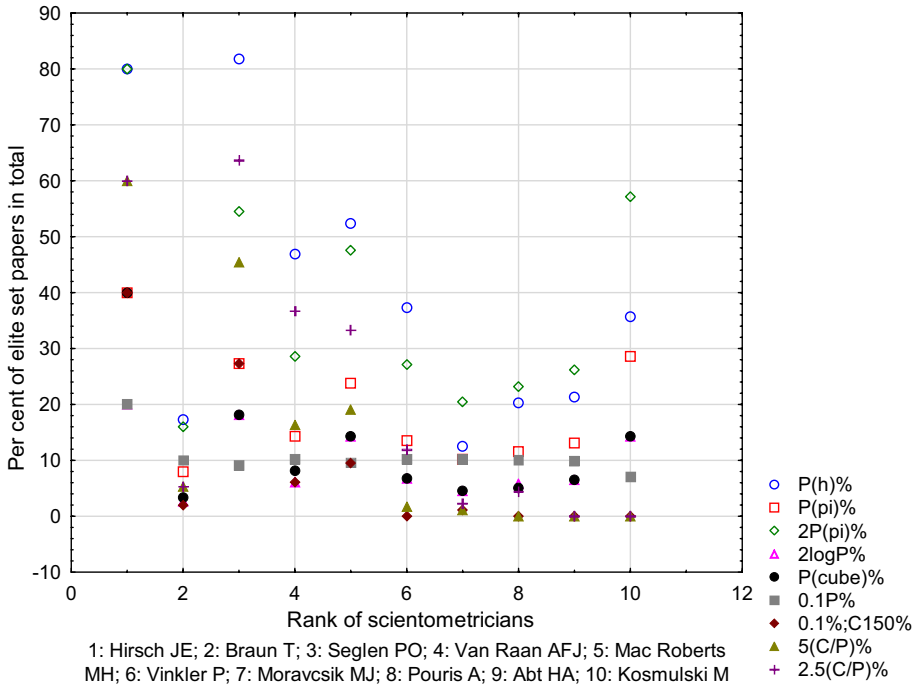

Fig. 1 Per cent of papers in different elite sets related to the total of the scientometricians studied. The decreasing number of citations in the *h*-core gives the rank of the scientometricians

Table 4 Spearman rank correlational coefficients between number of papers, P(S) in the scientometric part-set (see Table 1) and that in the corresponding elite sets, P(h), P(π), etc. (see Table 3)

	P(S)	P(h)	P(π)	2P(π)	2logP	$^3\sqrt{P}$	0.1P	$c_i \geq 150$	$c_i \geq 80$	$c_i \geq 40$
P(S)	1.00									
P(h)	0.73*	1.00								
P(π)	0.99*	0.75*	1.00							
2P(π)	0.99*	0.75*	1.00*	1.00						
2logP	0.93*	0.73*	0.94*	0.94*	1.00					
$^3\sqrt{P}$	0.92*	0.91*	0.94*	0.94*	0.89*	1.00				
0.1P	0.98*	0.74*	0.98*	0.98*	0.93*	0.94*	1.00			
0.01Pf $c_i \geq 150$	−0.15	0.14	−0.15	−0.15	−0.35	0.00	−0.10	1.00		
5Cf/Pf $c_i \geq 80$	−0.08	0.32	−0.05	−0.05	−0.23	0.13	−0.02	0.96*	1.00	
2.5Cf/Pf $c_i \geq 40$	0.05	0.59	0.08	0.08	0.00	0.31	0.12	0.74*	0.88*	1.00

The marked correlations are significant at $p < 0.05$

For obtaining *comparable* indices for characterizing impact of papers in elite subsets, the *number of citations to the papers in the elite subset* may be recommended. The number of citations in the elite subsets calculated by different methods is given in Table 5.

Figure 2 shows the number of citations in the different elite subsets by the scientists studied. The *rank number* of scientists (1, 2, 3, etc.) is the same as in Fig. 1. The data reveal, in most cases the number of citations is highest in P(h) or in 2P(π) set. The rank of the elite subsets by the number of citations strongly depends on the distribution of the publications by citation.

The number of citations (C) to the papers in the individual elite sets (h-core, π-core, 2logP-core, $^3\sqrt{P}$-core, etc.) correlated with the rank number of the scientists by the corresponding number of citations to the papers in the individual elite sets, yields negative exponential functions. Some of the equations are given as follows: $C(h) = 6949.547 \cdot e^{-0.390 r(h)}$ (see Fig. 3) and $C(\pi) = 5661.205 \cdot e^{-0.401 r(\pi)}$, where C(h) and C($\pi$) are the number of citations to *h*-core and π-core papers, respectively, whereas $r(h)$ and $r(\pi)$ are the rank number of scientists according to *h*-index and π-index, respectively. The corresponding equation for the citations (C) in the elite set obtained through the field standard $C \geq 5Cf/Pf$: $C = 7065.067 \cdot e^{-0.522 r(C \geq 5Cf/Pf)}$.

Spearman correlation coefficients (Table 6) between the numbers of citations in the individual elite sets reveal close correlation independent of the calculation method. In contrast to the correlations between the *number of papers* in the elite subsets (Table 4), in Table 6 all correlation coefficients ($r = 0.93$–1.00) between the *number of citations* obtained by *either inside or outside standard are significant* at $p \leq 0.05$. The correlation coefficient between the number of citations to the *h*-core and π-core is very strong: $r = 0.99$.

Model 2: Relating *h*-index and π-index of part-sets to *h*-index and π-index of the complex set

Table 1 reveals, h(S)-index of the part-set of the scientometricians containing exclusively scientometric papers, is *in all cases* lower than the h(E)-index of the complex set except for VanRaan (the value of both indices is 23). The difference between the indices may be

Table 5 Number of *citations* in different elite sets of the scientometric part-set (S) of publications of the studied scientists (see Table 1)

Name	Number of citations								
	Applied *inside* standards						Applied *field* standards		
	P(h)	P(π)	2P(π)	2logP	$\sqrt[3]{P}$	0.1P	0.01Pf $c_i \geq 150$	5Cf/Pf $c_i \geq 80$	2.5Cf/Pf $c_i \geq 40$
Abt HA	283	210	319	123	123	169	0	0	0
Braun T	2140	1683	2004	920	1036	1783	797	1244	1244
Hirsch JE	4120	3981	4120	3559	3981	3559	3981	4098	4098
Kosmulski M	56	50	69	33	33	20	0	0	0
MacRoberts MH	1129	933	1117	783	783	684	684	878	1022
Moravcsik MJ	629	607	679	478	478	607	337	337	408
Pouris A	368	254	385	160	160	233	0	0	160
Seglen PO	2120	1699	2028	1547	1547	1124	1699	1960	2070
VanRaan AFJ	1843	1087	1538	686	807	913	686	1171	1727
Vinkler P	867	491	745	304	304	405	0	114	453
Mean	1356	1100	1300	859	925	950	818	980	1118
SD	1234	1167	1205	1052	1173	1054	1235	1284	1271

P: total number of scientometric (S) papers (see Table 1)

Pf: number of papers in the field

Cf: number of citations obtained by Pf papers

extremely high (e.g. Hirsch: 4 vs. 58 or Seglen: 9 vs. 61) or lower (e.g. Moravcsik: 11 vs. 16). The π-index of the part-set of scientometric papers (S) is also lower in most cases than that for the corresponding complex (total, E = S + O) set. In two cases however, the indices are similar (MacRoberts: 9.33 vs. 9.33 and VanRaan: 10.87 vs. 10.87). The difference is extreme large for Hirsch (39.81 vs. 105.69) and for Seglen (16.99 vs. 103.54).

The very large differences in the impact indices may be attributed to the very high number of citations of the mentioned authors obtained to "other", non-scientometric papers.

The sum of π-index and h-index of part-sets: $H(S+O)$, $\Pi(S+O)$, respectively, i.e. the *cumulative* index is in all cases *higher* than the corresponding value of the complex set (Table 1). The ratios, $H(S+O)/h(E)$ however, may be different: e.g. Abt: = 42/30 = 1.40, whereas Seglen: 67/61 = 1.09.

The present study may be regarded as a first step for exploring relations of indicators calculated from data of complex sets containing part-sets. The number of the studied scientists is low, therefore most of the means in Table 1 do not significantly differ (at $p < 0.05$) from each other. However, the mean h-index for the complex sets, $h(E) = 30.30$ of the studied scientists significantly differs ($p = 0.01$) from the mean $h(S) = 13.80$ which refers to the scientometric part-sets. Nevertheless, the mean cumulative h-index, $H(S+O) = 37.20$ does not significantly differ ($p = 0.37$) from the mean h-index of the complex set, $h(E) = 30.30$. The relation between the mean cumulative π-index $\Pi(S+O) = 37.23$ to the corresponding index for the complex set, $\pi(E) = 31.52$ is similar. Although the difference of the mean $h(E)$ and $H(S+O)$ and $\pi(E)$ and $\Pi(S+O)$ referring to the set of individuals in Table 1 is not significant, on the *individual level* there
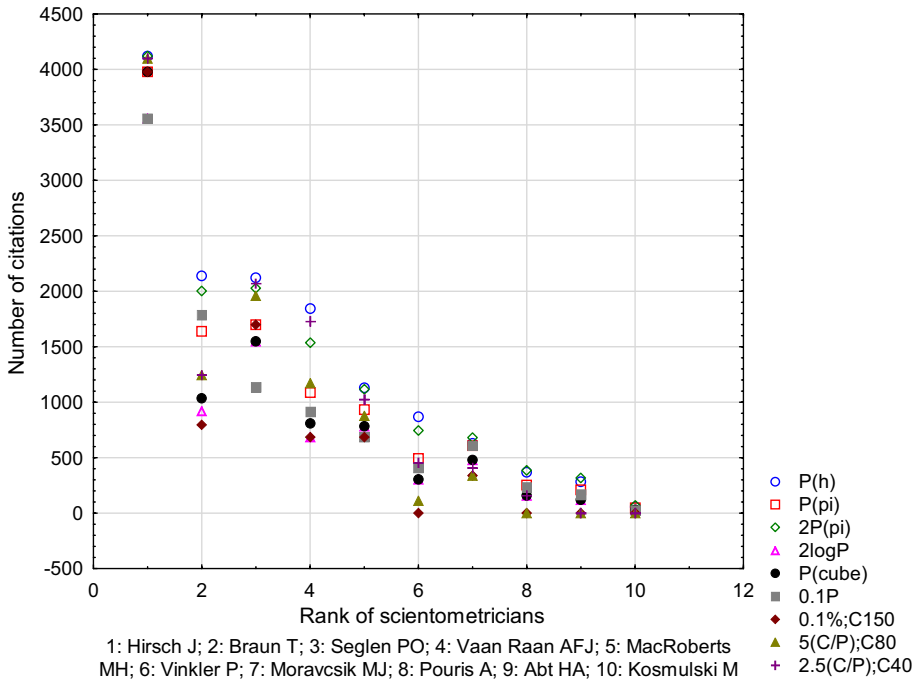
Fig. 2 Number of citations to the publications in the elite sets of journal papers of the studied scientists. The decreasing number of citations in the *h*-core gives the rank of the scientometricians

are great discrepancies. E.g. Abt: $h(E) = 30$ versus $H(S + O) = 42$ and similarly: Braun: $h(E) = 32$ versus $H(S + O) = 48$, or Moravcsik: $h(E) = 16$ versus $H(S + O) = 24$. The situation is similar also for the π-index. E.g. Hirsch: $\pi(E) = 105.62$ versus $\Pi(S + O) = 150.02$. This observation may indicate the application of the *cumulative index* (sum of indices obtained for the individual part-sets) for evaluating general (life) performance of individuals, instead of assessing impact of their publications by indicators obtained from the complex publication set containing publications from different fields.

The measure of difference between $h(E)$ and $H(S + O)$ or $\pi(E)$ and $\Pi(S + O)$ depends on several factors. E.g. number of papers in the different part-sets, difference in the bibliometric features of the corresponding fields, scientific level, relevance and timeliness of the papers, etc. The studied scientists are active on different fields outside scientometrics. This may be one of the reasons that the corresponding means of the total set of scientists do not significantly differ. Nevertheless, this observation should not ignore the selection of the publications of individuals according to fields. Surveying the $h(E)$ and $H(S + O)$ or $\pi(E)$ and $\Pi(S + O)$ data of the scientists listed in Table 1, it would indicate that selecting the papers in the complex sets of individuals and assessing them separately, would be highly relevant.

In order to clear some relations between the sum of h and π-indices of the part-sets and the h and π-index of the corresponding complex set, a model experiment was made.

Table 7 shows Model 2 with several examples for calculating h and π-index of part-sets with different number of papers derived from the corresponding complex sets. The question: *what is the relation between the h-index of the complEx set, h(E) and the sum*
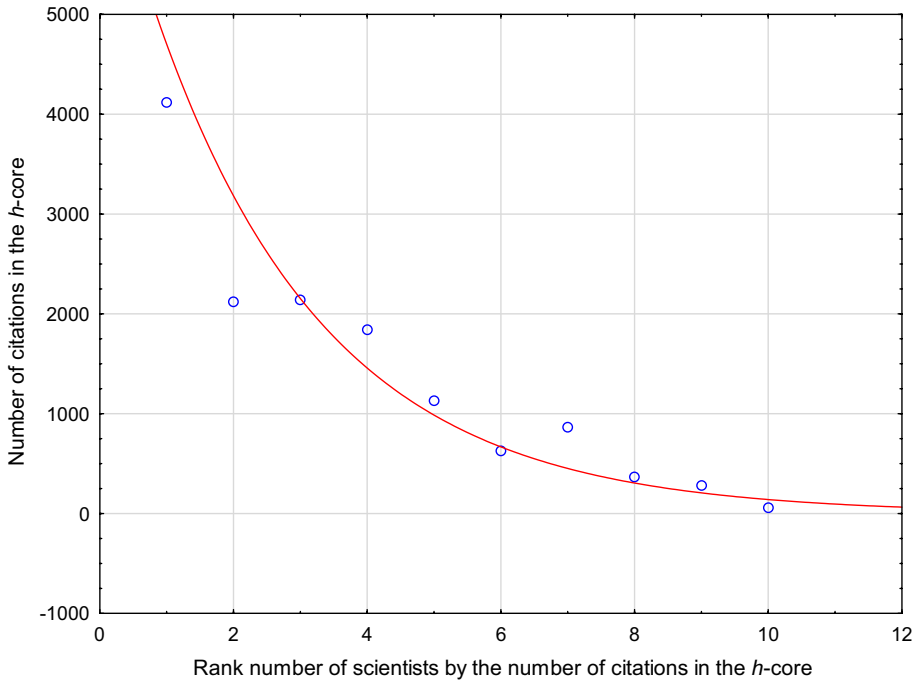
Fig. 3 Number of citations in the *h*-core of publications of the studied scientometricians correlated with the rank number of the scientists by the number of citations to the *h*-core papers

Table 6 Spearman rank correlational coefficients between the number of citations in the elite sets, P(*h*), P(*π*), etc. (see Table 5) of the corresponding scientometric part-set (see Table 1)

	P(*h*)	P(*π*)	2P(*π*)	2logP	$\sqrt[3]{P}$	0.1P	0.01Pf $c_i \geq 150$	5Cf/Pf $c_i \geq 80$	2.5Cf/Pf $c_i \geq 40$
P(*h*)	1.00								
P(*π*)	0.98*	1.00							
2P(*π*)	0.99*	0.99*	1.00						
2logP	0.96*	0.99*	0.98*	1.00					
$\sqrt[3]{P}$	0.98*	1.00*	0.99*	0.99*	1.00				
0.1P	0.99*	0.99*	0.98*	0.98*	0.99*	1.00			
0.01Pf $c_i \geq 150$	0.93*	0.97*	0.94*	0.96*	0.97*	0.96*	1.00		
5Cf/Pf $c_i \geq 80$	0.96*	0.99*	0.98*	0.98*	0.99*	0.98*	0.98*	1.00	
2.5Cf/Pf $c_i \geq 40$	0.96*	0.97*	0.98*	0.95*	0.97*	0.94*	0.93*	0.97*	1.00

The marked correlations are significant at $p < 0.05$

of h-indices of the corresponding part-sets (a1 and a2, b1 and b2...f1 and f2)? Similarly, what is the ratio between the sum of π-indices referring to the corresponding part-sets and the π-index of the complex sets?

Model 2 shows a publication set ("complex" set: E) consisting of 9 papers which received a total of 42 citations. The papers (1–9) are ranked according to decreasing number of citations. The h-index of the complex set is 5 because there are 5 papers in the set obtaining 5 or more citations. The π-index, $\pi = 0.01C(\sqrt{P})$ is 0.24 because there are $\sqrt{9} = 3$ papers in the π-core, and the number of citations obtained by the first three papers is equal to: $9 + 8 + 7 = 24$.

Case "A" in Table 7 represents two part-sets (a/1, a/2, differing from each other by research field) of which total number of papers is equal to that in the complex set: $E(P) = 5 + 4 = 9$. It is assumed that the number of citations to the individual papers in the corresponding part-sets is the same as that in the complex set. Table 7 shows, the value of h-index is equal to three for both a/1 and a/2, accordingly the sum: $H = 3 + 3 = 6$ is *higher* than the h-index of the complex set: $h(E) = 5$. The situation is similar for cases B and C.

It is obvious that the *sum* of h-index of two part-sets *cannot be lower* than the h-index of the parent complex set. Naturally, this relation is valid only if the *sum* of papers in the part-sets equals to the number of papers in the complex set. Further, the number of citations to the individual papers is the same in both the complex and in the corresponding part-set. If there is only a single paper in one of the part-sets with zero citation (see case f/1 in Table 7) – and this way the h-index of this part-set is equal to zero - the sum of indices of the part-sets: $H = (h1 + h2) = (0 + 5) = 5$ will be *equal to* the h-index of the complex set, $h(E) = 5$.

In cases D, F, and G the sum of h-indices of the part-sets *is equal* to the h-index calculated for the complex set.

Taking into account that the sum of h-indices of the part-sets can be *equal to or higher than* the corresponding index of the complex set, *six possible relations* can be assumed *between the h and π-indices*. Accordingly, the sum of the h-indices increases whereas the sum of π-indices also increases or decreases or it remains the same as the π-index of the complex set. Further, the sum of h-indices corresponds to that of the complex set, whereas the sum of π-indices decreases, increases or does not change related to the π-index of the complex set.

Model 2 demonstrates each of the mentioned opportunities. Case C, e.g. shows that the sum of h-indices of the part-sets ($h1 + h2 = 5 + 1 = 6$) is greater than the h-index of the complex set: $h(E) = 5$, whereas the sum of π-indices ($\pi1 + \pi2 = 0.17 + 0.03 = 0.20$) is lower than the value of the complex set: $\pi(E) = 0.24$.

Considering the data referring to A-G cases in Table 7, the conclusion may be drawn that the sum of h-indices of part sets can be *equal to or higher* than the h-index of the corresponding complex set. Further, the Model indicates that the sum of π-indices of the part-sets can be equal to or lower, and even *higher* that the π-index of the complex set.

Table 1 shows that the studied scientists belong to class A. Accordingly, both the sum of h-indices ($h1 + h2$) and π-indices ($\pi1 + \pi2$) of the part-sets (i.e. set of scientometric publications and set of papers on "other" topics) are higher than the h-index and π-index of the corresponding complex set.

Table 7 Model 2: Calculating h and π index and their sum for part-sets (a1, a2, b1, b2…f1, f2) derived from a common Parent Set (E) containing 9 papers

Rank of papers	Number of citations												
	Parent set (E)	A		B		C		D		F		G	
		Part set a1	Part set a2	Part set b1	Part set b2	Part set c1	Part set c2	Part set d1	Part set d2	Part set f1	Part set f2	Part set g1	Part set g2
1	**9**	9	7	9	5	9	2	9	8	0	9	9	5
2	**8**	8	6	8	2	8	1		7		8	8	1
3	**7**	4	5	7	1	7	0		6		7	7	0
4	**6**	2	0	6	0	6			5		6	6	
5	**5**	1		4		5			4		5	4	
6	**4**					4			2		4	2	
7	**2**								1		2		
8	**1**								0		1		
9	**0**												
Index													
P	**9**	5	4	5	4	6	3	1	8	1	8	6	3
C	**42**	24	18	34	8	39	3	9	33	0	42	36	6
√P	**3**	2	2	2	2	2	2	1	3	0	3	2	2
h(E)	**5**												
h1		3		4		5		1		0		4	
h2			3		2		1		4		5		1
H		6		6		6		5		5		5	
π(E)	**0.24**												
π1		0.17		0.17		0.17		0.09		0		0.17	

Table 7 (continued)

Rank of papers	Number of citations												
	Parent set (E)	A		B		C		D		F		G	
		Part set a1	Part set a2	Part set b1	Part set b2	Part set c1	Part set c2	Part set d1	Part set d2	Part set f1	Part set f2	Part set g1	Part set g2
$\pi 2$			0.13		0.07		0.03		0.21		0.24		0.06
Π		0.30		0.24		0.20		0.30		0.24		0.23	

P: total number of journal papers; C: total number of citations

$H = h1 + h2$

$\Pi = \pi 1 + \pi 2$

A: $h[a1] + h[a2] > h[E]$ and $\pi[a1] + \pi[a2] > \pi[E]$

B: $h[b1] + h[b2] > h[E]$ and $\pi[b1] + \pi[b2] = \pi[E]$

C: $h[c1] + h[c2] > h[E]$ and $\pi[c1] + \pi[c2] < \pi[E]$

D: $h[d1] + h[d2] = h[E]$ and $\pi[d1] + \pi[d2] > \pi[E]$

F: $h[e1] + h[e2] = h[E]$ and $\pi[f1] + \pi[f2] = \pi[E]$

G: $h[f1] + h[f2] = h[E]$ and $\pi[g1] + \pi[g2] < \pi[E]$

Model 3: Maximum value (Hmax) of the sum of *h*-indices (cumulative *h*-index) of two part-sets (*h*1 + *h*2) belonging to the same complex set

In studying the publication indicators of part-sets derived from a complex set, the question may arise: what is the maximum value of the sum of *h*-indices of the part-sets? Model 3 may yield an answer. The criteria of the Model given in Table 8 are the following:

- the complex set (E) contains 12 publications,
- each publication received 5 citations,
- the number of publications (12) in the complex set (E) is higher than two times the value of the *h*-index of the complex set: $2h(E) = 2 \cdot 5 = 10$,
- the sum of the number of papers in the part-set pairs (a1 and a2) or (b1 and b2) or…(g1 and g2) equals the number of papers in the complex set, i.e.: $P(E) = P(a1) + P(a2) = P(b1) + P(b2)…. = 12$.

Let's start with part-sets a1 and a2 containing 11 papers and a single paper, respectively. It is assumed that the number of papers gradually decreases in the first part-set (a1) from 11 to 6 (g1), whereas it increases in the second part-set from unity (a2) up to 6 (g2). The last two part-sets (g1 and g2) contain six papers each. The model shows that the sum of *h*-indices (H) of the part-sets increases from six up to 10.

Naturally, with more papers and citations in the parent set, the *h*-index of the parent set could change, and accordingly also the *h*-index of the part-sets and their sum (H) may increase. Nevertheless, the sum of *h*-index of the part-sets from the same parent (complex) set does not exceed two times the *h*-index of the complex set.

Figure 4 shows another example. In the presented model, a complex set may contain two part-sets (1 and 2) according to different topics of the papers. Let us investigate the change of the sum of *h*-index of the part-sets, dynamically. The rank number of the publication sets on the x-axis indicates the yearly condition of the sets. Let us assume that in the first year of the studied period part-set 2 contains 10 papers each cited 10 times, and these papers obtain no new citations in the next years. Accordingly the *h*-index of this part-set: $h2 = 10$ in each year. It is assumed further, there is only a single paper with a single citation in part-set 1, the first year. Accordingly the *h*-index of this set: $h1 = 1$. The sum of *h*-index of the part-sets in the first year: $H = h1 + h2 = 10 + 1 = 11$. In the second year the number of and citations to the papers in part-set 2 does not change. It is assumed however that in the second year part-set 1 will contain two papers each cited two times. Accordingly $h1 = 2$. The corresponding sum: $H = 10 + 2 = 12$. The number of papers in the corresponding *complex set* (part-set 1 and 2 together) is *12* in the second year, whereas its *h*-index remains 10, as there is 10 papers in the complex set cited 10 times each, and two papers with two citations each.

We may assume that the number of papers and citations would increase in set 1 gradually from year to year with *unity* up to 10. Accordingly, in the 10th year *the sum of h-index, i.e. cumulative h-index of the part-sets:* $(h1 + h2) = 20$. In the 10th year, the complex set (1 and 2 together) would contain 20 papers each cited 10 times. Consequently, the *h-index of the complex set is equal to* 10. Accordingly, the sum of *h*-index of the part-sets, i.e. *cumulative h-index is two times the h-index of the complex set.*

If in the next years the number of citations to the papers published in the first 10 years period does not change, whereas the possible new papers obtain 10 or less than 10 citations,

Table 8 Model 3: Maximum value of the sum of h-indices

r	E	a1	a2	b1	b2	c1	c2	d1	d2	f1	f2	g1	g2
1	5	5	5	5	5	5	5	5	5	5	5	5	5
2	5	5		5	5	5	5	5	5	5	5	5	5
3	5	5		5		5	5	5	5	5	5	5	5
4	5	5		5		5		5	5	5	5	5	5
5	5	5		5		5		5		5	5	5	5
6	5	5		5		5		5		5		5	5
7	5	5		5		5		5		5			
8	5	5		5		5		5					
9	5	5		5		5							
10	5	5		5									
11	5	5											
12	5												
Index													
P	12												
h(E)	5												
h1		5		5		5		5		5		5	
h2			1		2		3		4		5		5
H		6		7		8		9		10		10	

Relating the sum of the h-index (H=h1+h2) of part-sets (a1, a2, b1, b2, c1, c2, etc.) to the h-index of the parent set, h(E) with increasing number of papers (P=1, 2, 3, etc.) in one of the part-sets (a2, b2, c2, etc., respectively) whereas decreasing the papers in the other part-set (a1, b1, c1, etc.). The number of citations (c_i=5) to each paper is equal to the h-index of the parent set: h(E)=5, whereas the total number of papers (12) in the parent set is higher than the corresponding h(E) index

P: total number of papers in the parent set

h(E): Hirsch-index of the parent set

r: rank of publications

H=h1+h2=h(E)

The possible maximum value of the sum of h-indices of two part sets cannot exceed twice the Hirsch index of the parent set. Accordingly

Hmax=(h1+h2)max=2h(E)=10

the sum of h-index of the part-sets (1 and 2) would remain 20. If however the number of papers and citations increased in the complex set (and of course parallel in the part-sets), the h-index of the complex set and sum of h-index of the part-sets would increase accordingly. Nevertheless the highest value of the sum of h-index of part-sets within the corresponding complex set, could not surpass 2h(E), i.e.: Hmax=(h1+h2)max=2h(E).

Summary and conclusions

The evaluation of scientific publications of individuals is a difficult endeavour. The methods applied should correspond to the criteria of the assessment in question. Recently, methods based on scientometric indicators derived from only a subset (*core or elite set*) of the total are preferably applied. However, several scientists publish papers not only on a single
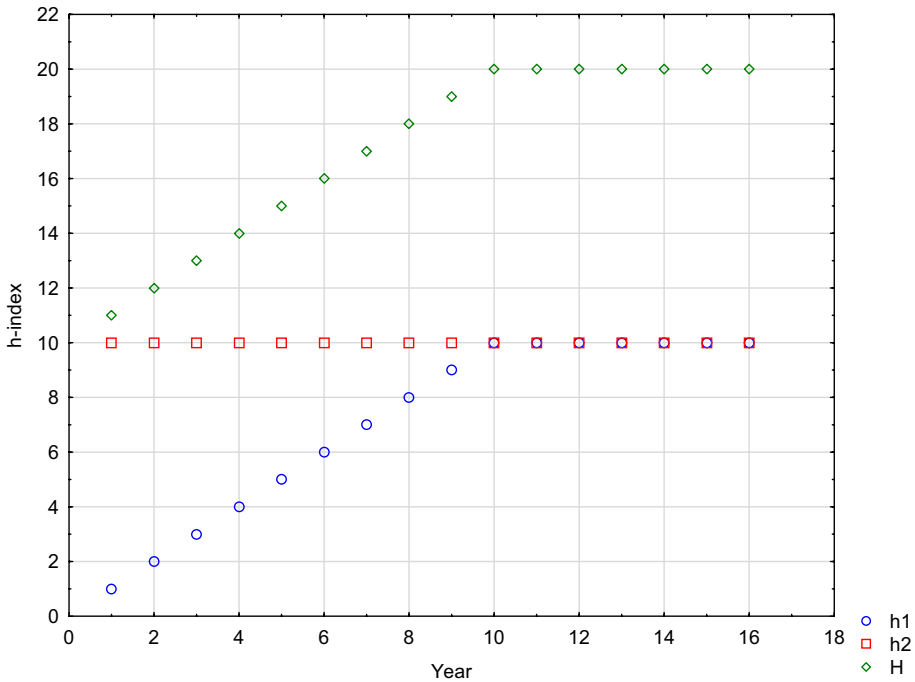
Fig. 4 Model 3: Possible maximum number (here: Hmax = 20) of the sum of *h*-index of two part-sets (here: *h*1 = 1, 2, 3…10 and *h*2 = 10, 10, 10…10). The part-sets contain 1 and 10, 2 and 10, 3 and 10…10 and 10 papers, respectively

but on several science fields or subfields of which bibliometric features (i.e. characteristic type of publications, developmental rate, citing or referencing norms and traditions, etc.) may be different. This phenomenon can be observed also by studying publications of authors in scientometrics. Therefore, in assessing publications of scientists it seems to be reasonable, *to separate their complex publication set into part-sets by fields* and, to calculate the corresponding scientometric indices accordingly.

From Model 2 presented, it concludes that the *sum* of *h*-index of part-sets within the corresponding parent (complex) set can be *equal to or higher than the h-index* referring to the corresponding complex (total) set. In contrast, the sum of π-index of the part-sets may be equal to or higher or even *lower* than the π-index of the corresponding complex set. The data of the publications of several scientometricians in Table 1 show that in practice, the *sum* of *h* or π index obtained for the part-sets separately, is greater in all cases than the same index calculated for the corresponding complex (total) set, as a whole. Accordingly $H(S+O) = h(S) + h(O) > h(E)$ and $\Pi(S+O) = \pi(S) + \pi(O) > \pi(E)$, where "S" refers to papers in the scientometric part-set, "O" to papers in the "other" (non scientometric) part-set, whereas "E" stands for all publications in the complex set, i.e. sum of S and O papers. The mentioned findings indicate that analysing complex publication sets of scientists would be reasonable only after selecting the publications according to the different scientific fields.

From Model 3 presented, it concludes that the possible *maximum value* of the *sum* of *h*-index of the corresponding part-sets originated from a common complex set is equal to *two times* the value of *h*-index of the complex (parent) set.

There are several statistics (e.g. *h*, *g*, π, percentage, etc.) offered in the literature (Vinkler 2017a, b) for calculating elite subsets. The *size* of the elite sets is not appropriate for assessing scientific eminence, except for the *h*-index. The *size* of elite sets strongly depends on the method selected. *Significant correlation coefficients* were obtained between the number of papers in the elite sets calculated by *inside standards* (i.e. standards derived from the total set analysed) (Table 4). This feature was found valid also for the number of papers in the elite sets calculated with *outside* (*field*) *standards*. In contrast, the correlation coefficients between the size of the elite sets obtained by *inside* standards and that obtained by *outside* standards *are not significant* (Table 4).

Relations concerning the *number of citations* in the elite sets obtained by different methods are different from that mentioned above (Table 5). Namely, the *correlation coefficients* between the number of *citations* in the different elite sets obtained by either *inside* or *outside* standard are *significant* (Table 6). This finding would indicate, scientists with high number of citations would show high impact according to any reasonable scientometric evaluation method.

The study may indicate some special consequences as to publication assessment of scientists. The evaluation exercises may refer to several aspects. They may aim to the *comparison* of the scientific performance of individuals active on a special field or topic. In selecting a person as university professor of Physical Chemistry e.g., we would not be primarily interested in the publications of the candidates on say, history, economics or library science. Accordingly, we should select the publications of the individuals *on physical chemistry*, and calculate the impact indices referring only to this field. It may happen however, that several scientists are suggested by a board to be decorated with a general (life-work) award (like "state prize" or "for merits in advancing science", etc.). In this case *all* scientific publications of the candidates independent of their topic should be taken into account. Accordingly, for characterizing *comprehensive* (*total*) *scientific performance* of scientists, we should use the *sum* (i.e. cumulative index) *of the corresponding indicators* calculated for the part-sets of their total, separately. Naturally, it would be relevant weighting the individual indices which refer to special fields. But, these aspects cannot be tackled in this paper.

The application of the *part-sets method* presented here may be maintained by the well-known fact that the bibliometric features depend on the field. Consequently, impact indicators, like *h*-index or π-index obtained from a mixed (complex) publication set (i.e. sets containing journal papers from several fields) cannot reflect the scientific merits of scientists correctly. For comparing scientific impact of researchers working on several fields, separating the publications according to the corresponding fields may be recommended.

It is obvious however that always the goals of the corresponding assessment should determine the possible items and methods applied.

References

Bornmann, L., Mutz, R., Hug, Sven. E., & Daniel, H.-D. (2011). A multilevel meta-analysis of studies reporting correlations between the *h* index and 37 different *h* index variants. *Journal of Informetrics, 5,* 346–359.

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

Iglesias, J. E., & Pecharromán, C. (2007). Scaling the *h*-index for different scientific ISI fields. *Scientometrics, 73,* 303–320.

Schreiber, M. (2010). Twenty Hirsch index variants and other indicators giving more or less preference to highly cited papers. *Annalen der Physik (Berlin), 52,* 536–554.

Schreiber, M., Malesios, C. C., & Psarakis, S. (2012). Exploratory factor analysis for the Hirsch index, 17 *h*-type variants, and some traditional bibliometric indicators. *Journal of Informetrics, 6,* 347–358.

Schubert, A., & Glänzel, W. (2007). A systematic analysis of Hirsch-type indices for journals. *Journal of Informetrics, 1,* 179–184.

Seglen, P. O. (1992). The skewness of science. *Journal of the American Society for Information Science, 43,* 628–638.

Todeschini, R., & Baccini, A. (2016). *Handbook of bibliometric indicators—Quantitative tool for studying and evaluating research* (pp. 1–512). Weinheim: Wiley.

Vinkler, P. (2007). Eminence of scientists in the light of the *h*-index and other scientometric indicators. *Journal of Information Science, 33,* 481–491.

Vinkler, P. (2009). The π-index: A new indicator for assessing scientific impact. *Journal of Information Science, 35,* 602–612.

Vinkler, P. (2010). *The evaluation of research by scientometric indicators* (pp. 1–313). Oxford: Chandos Publishing.

Vinkler, P. (2011). Application of the distribution of citations among publications in scientometric evaluations. *Journal of the American Society for Information Science and Technology, 62,* 1963–1978.

Vinkler, P. (2017a). The size and impact of the elite set of publications in scientometric assessments. *Scientometrics, 110,* 163–177.

Vinkler, P. (2017b). Core indicators and professional recognition of scientometricians. *Journal of the Association for Information Science and Technology, 68,* 234–242.

Vinkler, P. (2019). Core journals and elite subsets in scientometrics. *Scientometrics, 121,* 241–259.

Wildgaard, L., Schneider, J. W., & Larsen, B. (2014). A review of the characteristics of 108 author-level bibliometric indicators. *Scientometrics, 101,* 125–158.