 doi:10.20944/preprints202011.0164.v1

Does Impact Factor reflect whatsoever Scientific Importance of Publications?

Lev B. Klebanov[a]*, Yulia V. Kuvaeva[b], Zeev Volkovich[c †]

[a]Department of Probability and Mathematical Statistics, Charles University, Sokolovska 83, Praha 18675, Czech Republic

[b]Department of Finance, Money Circulation and Credit, Ural State University of Economics, Yekaterinburg 620144, Russia

[c]Software Engineering Department, ORT Braude College, 21982 Karmiel, Israel, e-mail: vlvolkov@braude.ac.il

This article is devoted to discussing the famous Impact factor citation suited to indicate the published scientific result's significance. Considering this problem based on the available statistical data, we demonstrate that: Impact factor follows a distribution having a heavy power tail that may lead to inconsistency in scientifically comparable journals' evaluation. Further, we remark that many papers recognized as sleeping and seemingly sleeping beauties are not considered when calculating. This makes the impact factor unsuitable in evaluating the original, groundbreaking discoveries and their consequences. Due to permanent changes in the composition of the underlying datasets, the impact factor can not be considered as a reliable indicator and to be used as a consistent evaluation metric. Overall, discussing all represented impact factor obstacles and other appropriate considerations, we conclude that the considered index can hardly be used as a reliable estimator of the publication's significance.

1. Introduction

In the 1950s, Eugene Garfield, the Institute for Scientific Information inventor, suggested the Impact Factor (IF) [3], which has probably been the most widely embraced bibliometric indicator exploited to evaluate scientific literature. Currently, this prevalent metric often signifies Journal Citation Reports *(JCR)*. Despite the known recognized strengths and limits, this metric stills commonly ill-treated and far exceeded the intentions for which has been aimed.

The IF look has started to be benefitted after publication in the Science Citation Index *(SCI)* [4]. At large, the idea is as such that the number of specific references to an article exposes the scientific interest and evaluates the average attention to this publication. The classical definition of IF, as given by Eugene Garfield, consists of a ratio attained as

†Corresponding author

*Partially supported by grant GAČR 19-04412S

dividing the number of citations of a source in the current year to the number of articles and reviews published here in the prior two years. JCR takes also into account as well as the published short notes. The matters being summarized in the denominator of the IF formulae are detectible in the Web of Science database as "Article," "Review," or "Proceedings Paper"[9]. A journal accepted for coverage in the Thomson Reuters citation database is evaluated by professionals counting the bibliographic and bibliometric appearances of all article categories. Items similar editorials, corrections, notes, retractions, and discussions are not considered. Numerous common problems associated with IF appearing in this area are summarized in [10]. However, it has become common to use the IF of a journal as an indicator of the scientific quality of the papers published in it. Mainly issues related to this practice are as follows [10]:

- Journal impact factors are not statistically representative for individual journal articles;
- Journal impact factors poorly correlate with actual citations of individual articles;
- Authors use many criteria other than IF when submitting to journals;
- Not all substantial publications are presented in the source database used for the indexing purpose;
- Books just recently have been included in the primary database;
- Journal set in the database vary from year to year;
- The citation rate of an article determines journal IF, but not vice versa.

Moreover, the limitations of the underlying database content profoundly influence the index, particularly by their continuous fluctuating from a year to another.

In this paper, we exhibit and discuss the weaknesses of the IF-like evaluation metrics partially resting upon mathematical model of the IF distribution proposed in Section 2. We demonstrate that the accepted IF applications are inherently not capable appropriately of reflecting and evaluating the scientific publications' genuine contribution. Such that, comparable journals may correspond to considerably different values of IF. Exploring situations caused by the behavior of the so-named "sleeping beauties" phenomena supports our inference and, together with an analysis of IF trends of several known journals. In the last part, we discuss the possible reasons for selecting the high IF journals suitable for potential submissions.

2. Heavy-tallness of IF and its consequences

As was mentioned earlier, IF of a journal in a given year y is a ration of the number of citations received in that year for the number of the articles published in the journal during the two preceding years:

$$IF_y = \frac{\text{Citations}_{y-1} + \text{Citations}_{y-2}}{\text{Publications}_{y-1} + \text{Publications}_{y-2}}.$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 November 2020 doi:10.20944/preprints202011.0164.v1

It is possible to relate it to any desired period. For example, the Journal Citation Report *(JCR)* also includes a five-year impact factor, which is similarly calculated within a five years phase [3].

The averaged value evaluates the citation of a separate article allegedly. This average is misused as an indicator of the quality of the published articles and, presumably, it should not be related to the number of articles issued in the journal. However, such expectations are not implemented according to the results given below.

Approaching the problem, we note that the number of papers published in a journal during a fixed time interval (the first source) and the number of citations for each publication (the second source) are, generally speaking, both random.

However, the first source of the randomness is almost absent in the journals issuing a fixed number of pages in each published volume. In the opposite case the role of randomness may be more essential. Therefore, we consider two separate cases corresponding to these publishing rules.

2.1. Nonrandom number of papers in an issue

Let $m \geq 1$ be the number of papers published in a journal for a period under consideration (say, 2 or 5 years). Generally, each of the papers gives rise to citations number. Many authors propose various distributions of a random variable X representing a number of a paper's citations (see, for example, [7]). It is commonly suggested that X's distribution has a heavy power tail with the infinite mean value. In other words, the tail possesses the following asymptotic property

$$\mathbb{P}\{X > x\} = \frac{A}{x^p} + o(x^{-p}) \quad \text{as} \quad x \to \infty. \tag{1}$$

Here $A > 0$ and $0 < p < 1$ are given constants and symbol $o(a)$ denotes the "little-o", i.e.: $o(a)/a \to 0$ as $a \to 0$. Empirical arguments supporting the presence of heavy tails in citation processes are given, say, in [8] and [2].

Aiming to understand the meaning of a long (heavy) tail of the citations' scattering intuitively, we must remind that the citations procedure may continue a very long period. It can be expressed out via a scientist's name or a recap of a related scientific event or object. For example, an allusion of the classical Archimedes' law in physics or the same Archimedes' principle in Mathematics, or another fundamental notion like the Euclidean geometry shows that this process may continue even more than 2000 years.

Let X_j $(j = 1, \ldots, m)$ be the numbers of citations of the j-th paper. We consider these variables as independent identically distributed random ones having probability distribution function F. Let us take the corresponding Laplace transform

$$L(s) = \int_0^\infty e^{-sx} dF(x).$$

The relation (1) shows that

$$\bar{F}(x) = 1 - F(x) = \frac{A}{x^p} + o(x^{-p}), \tag{2}$$

and

$$L(s) = 1 - Ap\Gamma(-p)s^p + o(s^p)$$

as $x \to \infty$ and $s \to 0$. Here $\Gamma(\cdot)$ is the Euler gamma function. It is not difficult to calculate that

$$\frac{1-L(s)}{s} = \int_0^\infty e^{-sx}\bar{F}(x)dx. \tag{3}$$

It is clear the IF is a random variable coinciding with the average

$$\bar{X}_m = (X_1 + \ldots + X_m)/m. \tag{4}$$

Because of independence and identically distribution property of the numbers of citations, the Laplace transform $L_m(s)$ of $\bar{X}_m$ has the following form

$$L_m(s) = L^m(s/m) \tag{5}$$

Similarly to (3), we have

$$\frac{1-L_m(s)}{s} = \int_0^\infty e^{-sx}\bar{F}_m(x)dx, \tag{6}$$

where F_m is the distribution function of $\bar{X}_m$.

Let us simplify the left hand side of (6) taking into account (eq1aa). We have as $s \to 0$:

$$\frac{1-L_m(s)}{s} = \frac{1-L^m(s/m)}{s} = \frac{1-(1-Ap\Gamma(-p)(s/m)^p + o((s/m)^p))^m}{s} \sim$$

$$\sim \frac{Ap\Gamma(-p)(s/m)^p m}{s} = \frac{Ap\Gamma(-p)m^{1-p}}{s^{1-p}}.$$

From here and (6) it follows

$$\int_0^\infty e^{-sx}\bar{F}_m(x)dx \sim \frac{Ap\Gamma(-p)\cdot m^{1-p}}{s^{1-p}}. \tag{7}$$

According to the known Tauberian Theorem, we get from (7) that as $x \to \infty$

$$\bar{F}_m(x) \sim \frac{Ap}{1-p}\cdot\frac{m^{1-p}}{x^p}. \tag{8}$$

Summarizing, we obtain the following theorem.

Theorem 1 *Suppose that the number of citations of a paper issued in a journal satisfies the relation (1), and a fixed number of papers in the interval under consideration is published. Then the journal IF in this period is a random variable with a distribution having a tail satisfying the relationship (8).*

Theorem 1 is genuinely essential for understanding the role of IF as a measure of a journal's scientific importance. First of all, the theorem shows that IF has a heavy tail under natural restrictions on the citation process verified in many studies (see, for example, [2] and [8]). Therefore, identical journals may have, from a statistical standpoint, dramatically different IF. Just this circumstance should prevent any attempt to employ IF as a measure of scientific importance.

Secondary, presence of a heavy tail implies that IF essentially depends on the number m of the published papers. This statement seems rather strange. Although, if the finite first moment (i.e. expected value) exits, then according to the law of large numbers, the ratio of the total number of citations to the total number of publications would be approximately equal to the mean value. More precisely, this ratio would converge to the mean value with an increase in the publications number. However, due to a heavy distribution tail, the first moment does not exist.

Let us assume that the distribution of IF belongs to the attraction domain of a one-sided stable law with the parameter $\alpha < 1$. Then the distribution of the variable

$$Y_m = \frac{\bar{X}_m\, m}{m^{\frac{1}{\alpha}}} = \bar{X}_m\, m^{\frac{1}{\alpha}-1},$$

where $\bar{X}_m$ is given in (4) converges to a one-sided stable distribution $\mathbb{S}$ with the parameter α.

However, $IF = \bar{X}_m$. Therefore,

$$IF = Y_m\, m^{1-\frac{1}{\alpha}}.$$

This relationship shows that the IF has a heavy tail induced by $\mathbb{S}$ and grows with the number m of the papers published in the journal. Therefore, one cannot compare different journals based on their IF without considering the numbers of papers published in each of the journals. Of course, this prevents from using journals IF to estimate the scientific value of papers published in them. Another consequence of a IF′s dependence on m is that IF , for a larger period, probably higher than for the shorter one. This fact is easy to see while comparing 2 and 5 years IF illustrated in the following figure

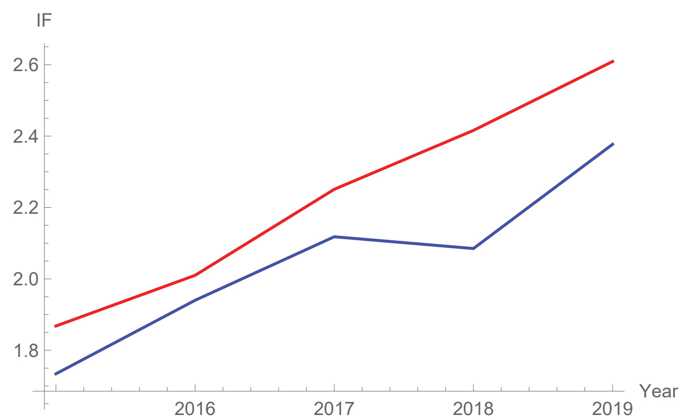

Figure 1. IF of the Annals of Probability during 2015-2019; the 5 years IF is in red, the 2 years IF is in blue.

An additional remark is that 2 (or just 5) years periods are often too short of revealing citations of essential innovative publications. To understand the genuine value of a new idea, one needs some time to study its connections with the previous results and to consider possible future perspectives. Therefore, the citations appearing in a short interval are mostly matching the papers elaborating on already proposed questions or previous ideas.

Let us study the relations between different types of $IF\prime s$ and consider Fig. 1, Fig. 2 and 3 containing a comparisons of two and five years $IF\prime s$.

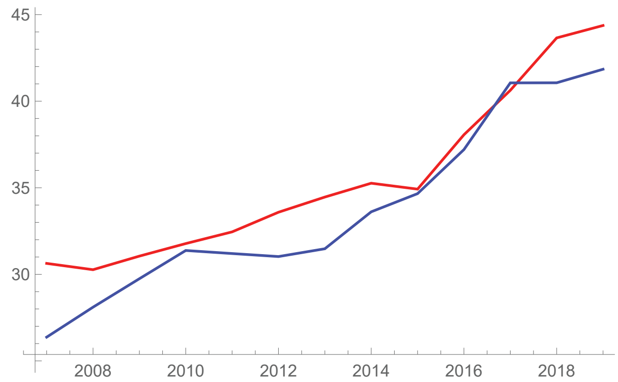

Figure 2. IF of the Science during 2015-2019; 5 years IF is in red, 2 years IF is in blue.

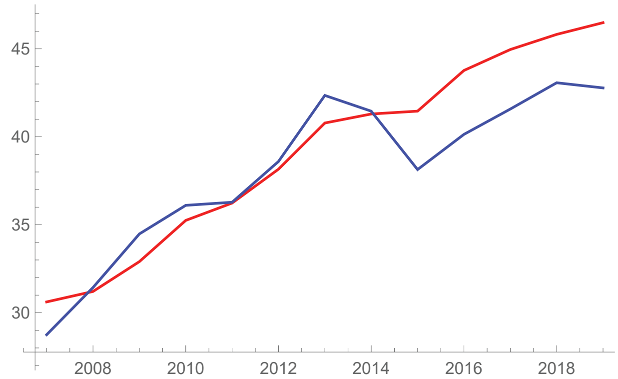

Figure 3. IF of the Nature during 2015-2019; 5 years IF is in red, 2 years IF is in blue.

If the first two charts clearly show that the 5-year index dominates, then in the last one, the tendency has changed in recent years. The results presented on Fig. 2 (Science) are similar to those in Fig. 1. However, the behavior of IF in Fig. 3 is different. This phenomenon can be partly affected by the exclusive role played by this magazine and its publishing creed. However, it seems that the more recently issued information is targeted

 doi:10.20944/preprints202011.0164.v1

not only at long-term global objectives but also in short period applications. Now, we consider the typical results given in Fig. 1 and Fig. 2. The aim is to justify more formally a reasonably evident fact that calculated over a more extended period, IF is expected to be higher.

Denote by $IF2_y$, $IF4_y$ and $IF5_y$ two, four and five years IFs calculated at the moment y. By c_y and κ_y denote the citations number and the publications number at the moment y, respectively. Then

$$IF2_{y+1} - IF4_{y+1} = \frac{c_{y-1} + c_y}{\kappa_{y-1} + \kappa_y} - \frac{c_{y-3} + c_{y-2} + c_{y-1} + c_y}{\kappa_{y-3} + \kappa_{y-2} + \kappa_{y-1} + \kappa_y} = \tag{9}$$

$$= \frac{(c_{y-1} + c_y)(\kappa_{y-3} + \kappa_{y-2}) - (c_{y-3} + c_{y-2})(\kappa_{y-1} + \kappa_y)}{(\kappa_{y-1} + \kappa_y)(\kappa_{y-3} + \kappa_{y-2} + \kappa_{y-1} + \kappa_y)}.$$

The sign of the difference $IF2_{y+1} - IF4_{y+1}$ is identical to the sign of the numerator in the right hand side of (9). However, the later is the same as the sign of $IF2_{y+1} - IF2_{y-1}$. It shows that if $IF2_y$ increases (decreases) in y with step 2 then $IF4_{y+1}$ is smaller (greater) than $IF2_{y+1}$. Of course, $IF5_{y+1}$ differs from $IF4_{y+1}$. However, if c_{y-4} and κ_{y-4} are small, then the difference between $IF5_{y+1}$ and $IF4_{y+1}$ is small and the difference between $IF2$ and $IF5$ may have the same sing as that between $IF2$ and $IF4$. These considerations show $IF2$, $IF4$ and $IF5$ are strictly dependent. Decreasing of $IF2$ between 2012 and 2015 explains further increasing of $IF5$ on the Fig. 3.

However, we see an additional drawback of IF. As it was mentioned above, the relation (9) shows that the increasing of $IF2$ with the time leads to the fact that $IF4 < IF2$ at the same moment. Of course, it is not the same that $IF5 < IT2$. However, the difference between $IF5$ and $IF4$ seems to be not too large. On the other hand, this difference may be large becouse of the dependence of IF on the number m of published papers.

2.2. Random number of papers in an issue

Let us now consider the situation with a random number of papers in an issue for the studied period and denote this random variable by M instead of the previously used constant m, aiming to stress our attention by supposing that M has the non-degenerate distribution. Introduce X_j as the random number of citations of a jth paper published in the journal under consideration, possing distribution function F and assume that these sequentially considered random variables compose a sequence of independent identically distributed random ones.

We are interested in the distribution of

$$IF = \frac{1}{M} \sum_{j=1}^{M} X_j. \tag{10}$$

The Laplace transform of IF given in (10) has the following form:

$$\mathbb{E} \exp\{-sIF\} = \sum_{k=1}^{\infty} L^k(s/k)\mathbb{P}\{M = k\}. \tag{11}$$

Let us consider two cases:

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 November 2020 doi:10.20944/preprints202011.0164.v1

1.
$$\sum_{k=1}^{\infty} k^{1-p}\, \mathbb{P}\{M=k\} < \infty, \tag{12}$$

2.
$$\sum_{k=1}^{\infty} k^{1-p}\, \mathbb{P}\{M=k\} = \infty.$$

In the first **case,** we have

$$\begin{aligned} L^k(s/k) &= (1 - Ap\Gamma(-p)\, s^p/k^p + o(s^p/k^p))^k = \\ &= 1 - Ap\Gamma(-p)\, s^p k^{1-p} + o(s^p k^{1-p}). \end{aligned}$$

Therefore,

$$\mathbb{E}\exp\{-sIF\} = 1 - Ap\Gamma(-p)\, s^p \sum_{k=1}^{\infty} k^{1-p}\, \mathbb{P}\{M=k\} + o(s^p). \tag{13}$$

However,

$$\mathbb{E}M^{1-p} = \sum_{k=1}^{\infty} k^{1-p}\, \mathbb{P}\{M=k\} > \sum_{k=1}^{\infty} \mathbb{P}\{M=k\} = 1$$

and all the consequences of the previous subsection are valued after changing m^{1-p} by $\mathbb{E}M^{1-p} < \infty$.

The distribution of IF has heavy tail and the probability of IF large values is higher for larger $\mathbb{E}M^{1-p}$. Instead of (8) we have as $x \to \infty$

$$\bar{F}_M(x) \sim \frac{Ap}{1-p} \cdot \frac{\mathbb{E}M^{1-p}}{x^p} \quad . \tag{14}$$

Consider now the second case**.** Let $F_M(x)$ be the probability distribution function of IF:

$$F_M(x) = \sum_{m=1}^{\infty} F_m(x) p_m,$$

where $p_m = \mathbb{P}\{M=m\}$ and $F_k(x)$ is the probability distribution function of $\bar{X}_m$ given in (4). For any positive integer K we have

$$\bar{F}_M(x) = \sum_{k=1}^{\infty} \bar{F}_k(x) p_k \geq \sum_{k=1}^{K} \bar{F}_k(x) p_k = \frac{Ap}{1-p} \sum_{k=1}^{K} k^{1-p}\, p_k \cdot \frac{1}{x^p} + o\left(\frac{1}{x^p}\right)$$

as $x \to \infty$. Therefore, as $K \to \infty$

$$\liminf_{x\to\infty} x^p \cdot \bar{F}_M(x) \geq \frac{Ap}{1-p} \sum_{k=1}^{K} k^{1-p}\, p_k \longrightarrow \infty \quad .$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 November 2020 doi:10.20944/preprints202011.0164.v1

It means that the tail of the distribution become more heavy than in the first case:

$$\lim_{x\to\infty} x^p \cdot \bar{F}_M(x) = \infty. \tag{15}$$

More detailed behavior of the tail depends on other characteristics of $F(x)$ that are not considered here becouse it is hardly applicable in practical situations.

Let us formulate the obtained result.

Theorem 2 *Let us suppose that the citation number of a paper published in the journal satisfies (1) and the number of the papers published within the interval under consideration is a random variable M. Then if (12) holds then the tail of the IF distribution satisfies (14); otherwise (15) is held.*

This theorem demonstrates that the statement of Theorem 1. stays be correct under suitable conditions also in the case when the number of articles in an issue is random, and the distribution of IF having has a heavy tail. Therefore, also here, journals , close in their nature, may possess dramatically different IF. This fact does not settle for us to consider the evaluation of publications quality with the help of IF as reliable and the index itself worthy of application.

3. IF, ”sleeping beauties” and “quasi sleeping beauties”

“Sleeping beauty” is associated with a scientific publication that remains overlooked (sleep) for a long time and then, practically unexpectedly, draws a great attentiveness (is awakened by a prince) (see, e.g. [11]).

Any scientist may provide some examples of such publications playing an essential role in the corresponding field. These articles cannot be taken into account while calculation 2 or 5 years IF. Indeed, IF ignores such an essential part of scientific literature that can be considered an additional obstacle to the index usage.

The traditional argument against this attitude is the fact that the sleeping beauties appeared very seldom. However, it is not so. There is given a precise definition in [6] of the sleeping period duration. As was shown, many papers had a rather long sleeping period and, therefore, were not accounted for correctly.

However, the sleeping period’s duration may generally play an essential role in truthful calculating the scientific impact of an ordinary paper. Namely, we do not suppose that a paper almost suddenly attracts a lot of attention. In fact, suppose that an article achieves a sequence of yearly citations $X_1, X_2, ..., X_j$, representing independent identically distributed random non-negative integer variables. Denote by $L(2)$ the first record moment, i.e.

$$L(2) = \min\{j : X_j > X_1\}.$$

It is known (see, for example, [1]) that

$$\mathbb{P}\{L(2) > j\} = \frac{1}{j}. \tag{16}$$

Therefore, the value $X_{L(2)}$ is not involved in the calculation of IF with the probability 1/2 and with the probability 1/5 in the calculation of the 5 years IF. The situation appears similar to one arising for the sleeping beauties but in entirely different circumstances. A random character of the citations number make it possible to have a large number of citations occur after a relatively long period with just small numbers of citations. It seems like the "sleeping beauty" situation; we call such case "quasi sleeping beauty". The appearance of many articles citing this publication could lead to a purely psychological reaction. This event also manifests itself in a continuing increase in the citations.

The described situation of "quasi sleeping beauties" demonstrates an appearance of random behavior of the citation numbers similar to that for "sleeping beauties", but without a significant essential scientific value. The IF fashion does not allow either to comprehend the real reason for such behavior nor just notice it.

Some examples of the inability to distinguish between these options, even for highly qualified experts, are known. In this regard, it may be mentioned a well-known event of a negative review of Michail Ostrogradskij to the fundamental Nikolay Lobachevsky's paper setting forth the foundations of an innovative non-Euclidean geometry (see, [**?**], p.p. 252-255).

To use the Sibuya law as a citation distribution is proposed in [7]. Fig. 4.presents a simulated sample of 75 items from this distribution.

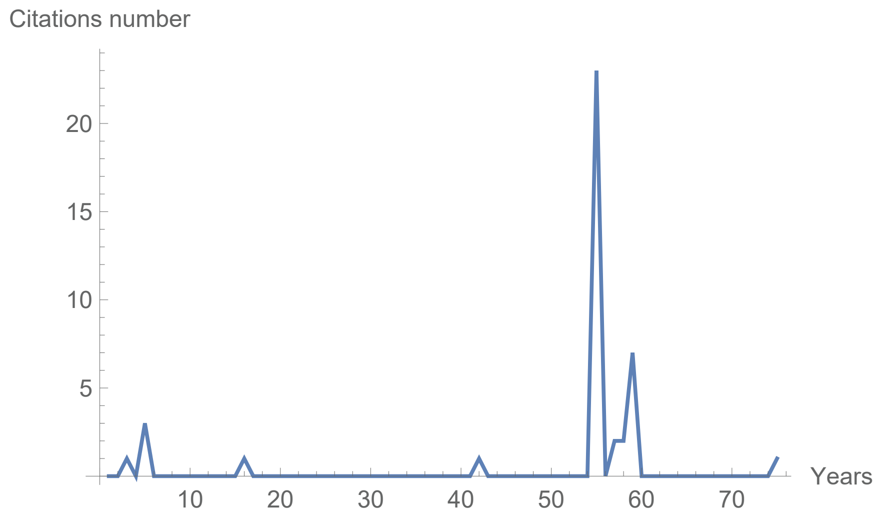

Figure 4. Simulated random citation.

We consider each observation as a citations number in the corresponding year (total, 75 years). For the first two years, there are no citations at all. However, after 54 years, the number of citations becomes larger. Of course, it is not a "sleeping beauty", but one can consider (by mistake!) this fact as a sign of growing interest in this publication; however, it is just a simulation of a typical situation appearing with a paper. Thus, each article can be comprehended as such a publication, and its destiny depends on many circumstances of an accidental nature.

4. Causes of nonrealistic citations' trends reflected by IF

From the proposed model's standpoint, the situation trend's behavior appears to be very strange overall. We illustrate this topic on an example of the "Nature" and "Science" journals between 1997 and 2019 years. The data are given in Fig. 5. An upward trend for both journals is clearly observed. For "Nature", it starts in 2005, and for "Science" in 2007.

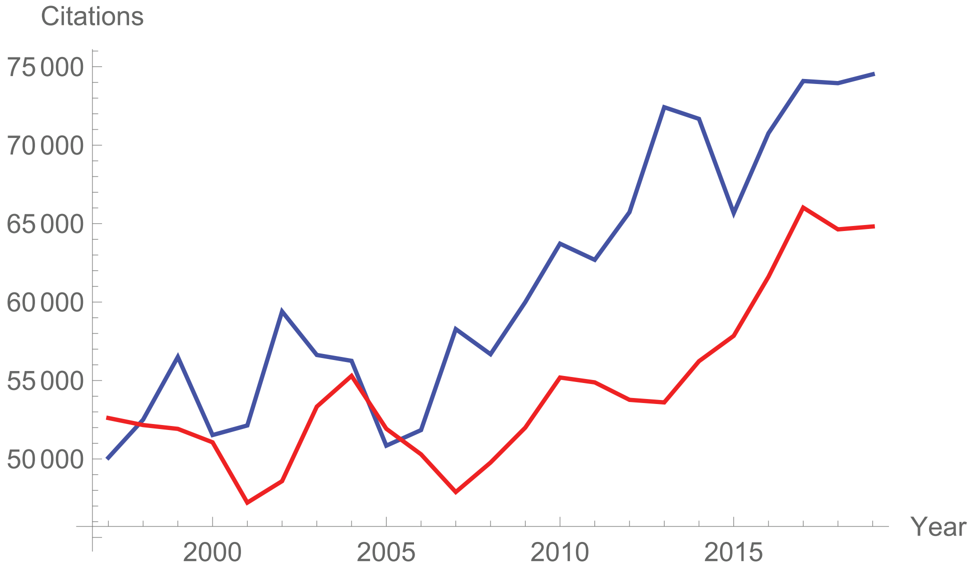

Figure 5. Graphs of Nature (in blue) and Science (in red) citations

A similar situation happens for the case of IF (see Fig. 6). Such a rapid growth seems to be not realistic, mainly because of the published results' multidisciplinarity, naturally leading to many possible sleeping beauties (see, [6]).

However, how can one explain the observed data? The most likely reason for this is a change in the underlying databases happened in reality. From some changepoint, books have been partially, only those published by some chosen publishers, taken into account. A similar situation appears in articles where the list of the perused journals has been expanded over time. Of course, these alterations are unlikely to improve $IF\prime s$ ability to reflect the real scientific significance of publications. Moreover, it makes the indices' values calculated for different lifetimes hardly comparable.

It is essential to mention that similar events occur not only for such notable journals as "Nature" and "Science", playing the role of a representative example in our considerations. However, a growth configuration may depend on the field of science and the corresponding employed database.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 November 2020 doi:10.20944/preprints202011.0164.v1

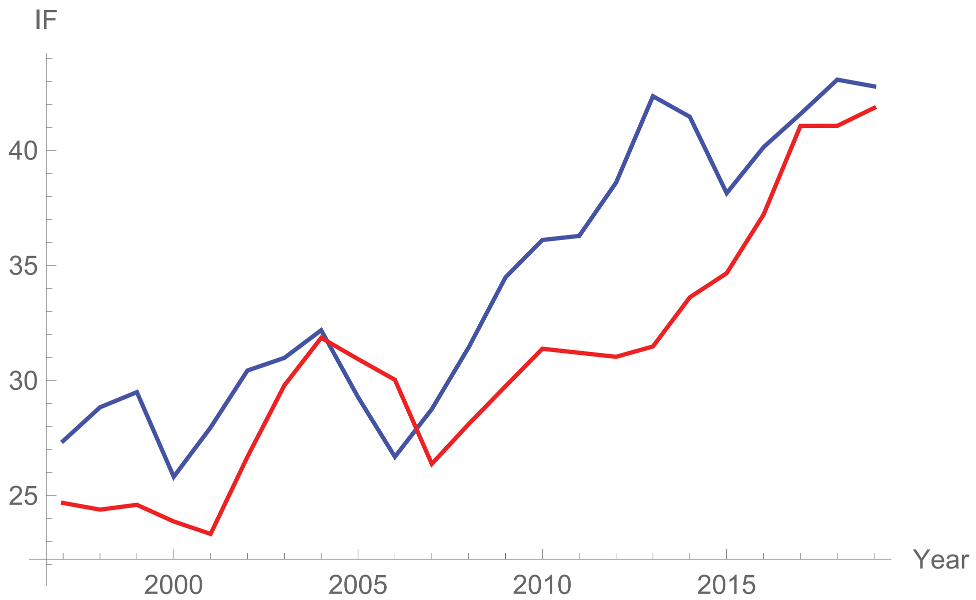

Figure 6. Graphs of Nature (in blue) and Science (in red) citations

5. Conclusion

As well known, one of the reasons for a choice of an appropriate journal for a possible submission is the journal's scientific reputation reflected by its IF. The high-level journals such as "Nature", "Science", and "PNAS" are acknowledged to provide a genuinely high level of the peer review process (although not without annoying exceptions) and are considered overall preferred.

The motivation for selecting a more dedicated journal with high (but not the highest) IF is based on many often contradictory considerations. A high value of IF is merely psychologically associated with a high quality of the journal. Another important consideration is the ability to publish a paper reasonably quickly. Nevertheless, the time needed to assess the fairness and significance of an innovative thought or a non-standard approach can be reasonably longer. Of course, there is a contradiction between the desire to publish a paper rather quickly and to have a high scientific publication index's score.

The issuing process of an article meets many serious problems. It is difficult to find a large number of highly qualified reviewers. Young reviewers tend to be too radical to appreciate significantly different views from their own. More mature scientists incline to overestimate the role and the importance of the traditional concepts and may also misjudge new ideas. Besides, they are usually swamped and do not have sufficient means to deal with a review. The anonymous reviewers' practice may reduce their responsibility for a review quality. These circumstances can often lead to refusals to take part in the review process covered up with various formal excuses usually presented without further explanation.

Of course, such reasons are more common for journal publishing resting upon commercial success. The assessment of scientific quality based on the reviews' conclusion looks more than strange, although it is very acceptable. In our opinion, even the presence of a negative review (or several negative reviews) is not yet a strong enough reason to reject a manuscript (of course, if the author himself disagrees with the reviewers' opinion). It seems that it is more advisable to publish both the manuscript itself and the signed reviews so that readers can judge the paper himself: "truth is born in a dispute".

During the review process, two types of errors commonly appear. The first one is the publication of a low-quality article, and the second one is the rejection of a significant new result. It seems that it is much more essential to prevent the unjustified rejection of a significant article.

An article's quality assessment based just on the journal IF leads to young scientists' motivated orientation towards publications only in such journals. However, as we have shown in this article, the IF index hardly expresses the scientific value of a publication. In other words, there is a substitution of scientific attention by commercial success.

The following shortcomings of IF are analyzed in this paper:

- A heavy tail of the IF distribution;
- Dependency on a journal capacity;
- Existence of a large number of nonseen “sleeping” or “seemingly sleeping” beauties;
- Alternation of the underlying databases;
- Self-seeking principles of selecting articles in some journals with high IF.

These results evidence that IF can hardly serve as an indicator of the scientific significance of the scientific articles.

REFERENCES

1. M. Ahsanullah and V.B. Nevzorov, Records via probability theory, Atlantis Press (2015).
2. E. Callaway, Beat it, impact factor! Publishing elite turns against controversial metric, Nature, 535 (2016) 210.
3. E. Garfield, Citation Indexes for Science: A New Dimension in Documentation through Association of Ideas, Science, 122 (1955) 108.
4. E. Garfield, "Science Citation Index"A New Dimension in Indexing, Science, 144 (1964) 649.
5. V.F. Kagan, Lobachevsky, Academy of Sciences of USSR, Moscow-Leningrad, (1948) (in Russian).
6. Q. Ke, E. Ferrara, F. Radicchi, and A. Flammini, Defining and identifying Sleeping Beauties in science, PNAS, 112 (2015) 7426.
7. L.B. Klebanov, Yu.V. Kuvaeva and Z.E. Volkovich, Statistical Indicators of the Scientific Publications Importance: A Stochastic Model and Critical Look, Mathematics, 8 (2020) 713.

8. V. Larivière, V. Kiermer, C.J. MacCallum, M. McNutt, M. Patterson, B. Pulverer, S. Swaminathan, S. Taylor, and S. Curry, A simple proposal for the publication of journal citation distributions, BioRxiv (2016) 1.
9. M.E. McVeigh and S.J. Mann, The Journal Impact Factor Denominator: Defining Citable (Counted) Items, JAMA, 302 (2009) 1107.
10. P.O. Seglen, Why the impact factor of journals should not be used for evaluating research, BMJ, 314 (1997) 497.
11. A.F.J. van Raan, Sleeping Beauties in science, Scientometrics, 59 (2004) 467.