THE CITATIONS CHASE: A RANDOM RESEARCHER'S VIEW

Harilaos N. Psaraftis
Department of Technology, Management, and Economics
Technical University of Denmark
2800 Kgs Lyngby, Denmark

ABSTRACT

The purpose of this paper is to express some thoughts on what I term "the citations chase". By this I mean the importance attached to citations as an indicator of academic excellence, prominence, impact, influence, and the like. Bibliometrics is a term used to describe the science of citations, and activity in this area has grown considerably over the years. The paper has been triggered by some recent bibliometric studies in which I was lucky to find my name being included. Even though the exposition in this paper is, to some extent, subjective and qualitative, an attempt is also made to also address some quantitative aspects of the subject, which may provide some additional insights on what may lie below the surface of citations statistics. To that effect, and among some higher order citations statistics, the "claim to fame ratio" is defined. Some rudimentary analysis of a 100,000+ scientists database focuses on Logistics & Transportation and on Operations Research.

1. Introduction

The purpose of this paper is to express some thoughts on what I term "the citations chase". By this I mean the importance attached to citations as an indicator of academic excellence, prominence, impact, influence, and the like. Bibliometrics is a term used to describe the science of citations, and activity in this area has grown considerably over the years. There are citations statistics, more statistics and even more statistics produced, analyzed and discussed, sometimes ad nauseam. The dictum "publish or perish" has transcended into "be cited or perish". Just publishing is not enough.

Triggers for this paper have been some recent bibliometric studies in which I was lucky (and surprised I must say) to find my name being included. These include (in reverse chronological order):

a. A bibliometric study for the last 20 years of the *Networks* journal, as described in Golden and Shier (2020).
b. A bibliometric overview of the last 40 years of the *Transportation Research Part B: Methodological* journal, as described in Jiang et al. (2020).
c. A bibliometric overview of the *WMU Journal of Maritime Affairs* journal since its inception in 2002, as described in Sahoo and Schönborn (2020).
d. A 100,000+ "highly cited researchers" database, as described in Ioannidis et al. (2019).
e. An analysis of the top 50 authors, affiliations and countries in maritime research, as reported in Chang et al. (2018).

First of all, a clarification is in order. It feels nice if one's paper is cited, and it may feel awful if no one bothers about it. Also, and by and large, no one can claim to be a serious academic if he or she cannot display a reasonable number of citations for his or her work.

At the same time, and having been in academia for close to 42 years and having studied the subject as best I could (although I admit non-exhaustively), I honestly believe that, even though I can understand the importance of citations, I also feel that this matter is overblown out of proportion and really distracts from more important things in life, academic and other. It does not do justice to many people, including those who consider citations as extremely important. Further, sometimes there is more than meets the eye if one is to dig deeper in the citations statistics.

Even though the exposition in this paper is, to some extent, subjective and qualitative, an attempt is also made to also address some quantitative aspects of the subject, which may provide some additional insights on what may lie below the surface of citations statistics.

The rest of this paper is organized as follows. Section 2 goes over some citations basics. Section 3 comments on the 100,000+ scientists citations database, with a focus on Logistics & Transportation and on Operations Research, and conducts a rudimentary analysis that differentiates citations statistics based on the population of citations versus the population of papers. To that effect, and among some higher-order citations statistics, the "claim to fame ratio" is defined. Section 4 makes some concluding remarks.

2. Citations basics

Citations count how many times a particular paper is cited (or referenced) by other papers. Whenever a paper is cited, all authors of the paper get equal credit for the citation, even though some citations engines also keep track of who is the first author, who is the last author and obviously if one is the sole author. In some disciplines (but not in all) sometimes the last author is considered as an honorary position. Depending on how the contribution among co-authors is shared, the order of appearance of the co-authors is determined by themselves, however there are no unambiguous rules. In an equal-contribution paper, order can be typically alphabetical, even though this may be unfair to authors closer to A and does not favor authors closer to Z. They might as well toss a coin. I am lucky that in the Latin alphabet my last name begins with a P, versus a Ψ in the Greek alphabet (one before last letter Ω). By the way, Z is the 6^{th} letter in the Greek alphabet. If contributions are not equal, authors may be displayed by decreasing share of contribution. But there is really no easy way for an external person to assess or verify contributions, even though recently publishers ask authors to write a little blurb specifying this kind of information. Whatever it is, if a paper has N authors, each of the N authors gets one citation each time the paper is cited by another paper.

According to this system, a sole author who gets one citation for a paper, and an author of a paper with $N>1$ authors that also gets one citation are equal in terms of citations. Each of the N co-authors of the N-author paper will get one citation, the same as the citations of the sole author paper. This, in my opinion, defies any reasonable sense of fairness, especially if N is high. Cases in which N is between 2 and 5 are very common, but there are also cases in which N is much higher. N can exceed 5,000, see later example.

It is clear that the system can be abused. Assume that we hypothetically have 10 distinct papers, each written by a distinct sole author, and all these papers are published and each is cited once, so each author will get one citation. But if all 10 authors collude to co-author the same 10 papers and again each paper is cited once, each author will get 10 citations. What has changed? Only the number of authors. This is of course an extreme example, but in general more authors favor more citations.

Of course, if citations engines counted citations differently, for instance by assigning a *fractional citation* of 1/N to each of the N authors of a cited paper, as it would seem the fair thing to do (at least as a first order approximation), things would be very different. I suspect we would have a completely different picture as regards what papers are written by whom, let alone who is highly cited and who is not. But I consider the likelihood of this happening very low.

The citing papers may be authored by other authors or by some of the authors of the cited paper. If one cites his or her own paper, it is a self-citation, and it is typical to exclude self-citations in any serious citations count. Citations engines such as Web of Science or Scopus detect self-citations. There exist pathological cases of some authors getting most of their citations from themselves.

For what things are citations important? They seem to be important in hiring, promotion, and tenure decisions, not to mention salary negotiations and other things, such as funding proposals and the like. As mentioned earlier, they are considered a proxy for academic excellence, prominence, impact, influence, and the like. Certainly they are not the only proxy, but they are becoming increasingly important.

According to Google Scholar, which is one of the citations engines, the highest cited scientist is Michel Foucault, with more than 780,000 citations. Also, a high number of citations in a paper usually correlates with journal prestige. As an example, a 1999 paper in *Science* on the emergence of scaling in random networks (Barabasi and Albert, 1999) has more than 37,000 Google Scholar citations (it is interesting that the above paper was submitted on June 24, 1999 and accepted on Sep. 1, 1999).

When I was a faculty member at MIT (1979 to 1989), I do not recall citations playing a major role in my promotion to associate professor (1983) or to the tenure I received (1985). At least I was not explicitly aware of it. I was aware of the *Science Citations Index*, but that was about it. My guess is that citations may have played an implicit role in these decisions, but I recall that I was more preoccupied with the number of papers I had written and publishing them in good journals than counting their citations, which is something I did not do myself. Maybe someone else did this for me, but it was not involved. In retrospect, academic life was much simpler then, even though these evaluation processes were very elaborate.

The first time I seriously noticed citations was in the context of some faculty promotions at the National Technical University of Athens (NTUA) in Greece. And then I started looking at my own citations, which could be found online. One thing that surprised me was that even though I continued to publish, most of the citations I was getting were attributed to papers I had written while at MIT. I do not think this was due to the switch of affiliation, but I suspect it was due to the switch of the topics I was mainly writing papers on. At MIT most of my papers were in Operations Research, while at NTUA I basically switched gears and worked on things like Maritime Safety, Maritime Economics, Ship Emissions, and all kinds of diverse other topics. Operations Research ceased to be a main priority for me, even though I attempted to partially reengage in it during the last 10 years or so. I observed that papers in the new areas I published received far fewer citations, without knowing why. But this did not really bother me.

Then at some point came the 5 ½ year period (1996 to 2002) when I was CEO of the port of Piraeus, a major port in the Mediterranean. I wrote *zero* papers during that period, even though I sent an article to *OR/MS Today* (Psaraftis, 1998). This whole period, which I consider as very important in my whole professional life, counted very little or nothing, citations-wise. I wrote a couple of papers about the port after I left, which attracted very few citations. I resumed my research activities afterwards, but at no point I was preoccupied with citations. Getting funding for my projects seemed like a much more important priority.

3. The 100,000+ scientists database

It was as late as fall 2019 that I became aware of what some people refer to as the "100,000 highly cited scientists" database, which is described in Ioannidis et al. (2019). The actual number of people listed in it is 105,000. A colleague told me that I was in that database, which to me was a pleasant surprise. Still, I did not pay too much attention on who constructed the database, how the database was constructed, and who else was in it.

My interest in that database resumed in 2020 with all the publicity Professor John Ioannidis of Stanford University (an epidemiologist) received in the context of his views on COVID-19 (see for instance, Ioannidis (2020)). I did not know about him until then and I discovered that he had an exceptionally high number of citations, something like 30 times mine. I then discovered that he was the main architect behind this database. As many as 40 indices were collected for each of the scientists in the database and people were ranked according to a composite index that combines 6 citations metrics. To give a bit of perspective, Ioannidis is No. 52 on this database, and I am No. 53,647.

Reading the (very interesting) paper that describes the database, I discovered that some 6,880,389 scientists who had published 5 or more papers formed the basis of that database. Of those 6,880,389, the top 100,000 (in fact 105,000) were selected and ranked. The database breaks down scientific disciplines into 22 scientific fields and 176 subfields, according to the Science-Metrix journal classification system (Archambault et al., 2011). Fields include broad areas such as Engineering, Economics and Business, Biology, Chemistry and others (22 total), and the subfields go into further detail, for instance Optoelectronics & Photonics, Urban & Regional Planning, Distributed Computing, Operations Research, Logistics & Transportation, and many others (176 total). In that sense, a scientist can be ranked not only within the entire database, but also within the field and subfields that are associated with the scientist. To that effect, and for each scientist, the database provides the most common scientific field and the two most common scientific subfields of his/her publications, one main and one secondary. By doing so, one is supposed to be able to compare scientists that work in the same areas, as it would make little sense to compare an astronomer with a psychologist, or an economist with a nuclear physicist. Thus, the database may allow comparisons and rankings of people working on the same subjects.

Looking at the database and accompanying paper, I was first impressed by the fact that people in some areas get widely differing citations numbers than people in other areas. For instance, the 99-percentile of citations in Nuclear and Particle Physics is 32,708 (meaning that the percentage of scientists in that subfield who have citations up to 32,708 is 99%). But in Astronomy and Astrophysics the 99-percentile is 16,244, in Ecology it is 8,302 and in Cultural Studies it is only 588! I found these numbers, along with the populations of each field and subfield, to be literally "all over the place". For some subfields like Drama & Theater, or Music, there are zero people in the top 100,000+ list. But other areas, like for instance Applied Physics, are very well represented in that list. This means that for some disciplines it may be easier to get into the 100,000 list than for some other disciplines.

Then the next question that came to mind was, why do people in Nuclear and Particle Physics get so many more citations than people in Ecology? Is it because these folks have higher IQ, work harder, or are just more prolific? Is it because it is easier to publish in that area? Is it because work can be dissected into narrower subjects, each being the subject of a paper? (for instance, prove a theorem in mathematics, or prove a property of a sub-atomic particle). Is it because there more interest in some areas than in some other areas? Or is it because of another reason? To be honest I do not know. I suspect that the number of authors must also play some role, and papers in Nuclear and Particle Physics typically have many authors. The record is a paper with 5,154 authors, published in *Physical Review Letters*. Only the first 9 pages of this 33-page paper (Aad et al., 2015) describe the research itself, including the references. The rest of the paper lists the authors.

Judging from myself, more citations do not necessarily imply author preference or favor of a paper. The favorite paper that I wrote, Psaraftis (1984), published in *Networks*, has very few citations. I like it as much as I did when I wrote it, in spite of the few citations.

Also, getting more citations does not necessarily mean that people agree with a paper's results, or that they even consider the paper important. One may get many citations if a lot of people disagree with one's paper, or if people say that the paper is nonsense. Ioannidis's perhaps most famous paper was a paper in which he proved that most of published research findings are false (Ioannidis, 2015). That this paper is being highly cited (more than 9,300 Google Scholar citations) does not necessarily mean that those who cite it also agree with it. This would amount to their consent that their own work is probably false. There is no such thing as a negative citation.

The other thing I noticed, with some surprise, is that the 100,000+ database omits some people I know and who I thought would be included. Whether this is due to errors in the database, or these folks not making the 100,000 person list, I do not know. After all, if the database is wrong, it is a proof of Ioannidis's own theory.

Digging deeper into the database, I thought that some interesting "higher order" citations statistics can be defined. Suppose a scientist has X citations. Then we may want to know the following:

(a) What is the value of C_s, defined as the percentage of X within the population *of that person's citations*, that is ascribed to papers in which he or she is sole author?

(b) What is the value of P_s, defined as the percentage of X within the population *of that person's papers*, that is ascribed to papers in which he or she is sole author?

(c) How do the values of C_s and P_s compare?

Percentages defined by (a) and (b) can be thought of as *second-order* citations statistics, being ratios of first-order citations statistics. They can also be defined *in a broader sense* as C_{sf} (or C_{sfl}), that is, percentages of X within the population of that person's citations, that are ascribed to papers in which he or she is sole or first author (or sole, first or last author). Similar definitions pertain to P_{sf} (or P_{sfl}).

Why would we want to know these percentages? Because it may be interesting to know the extent of a particular author's own contribution to that person's collection of papers, and also to his or her citations. Irrespective of whether or not an author is prolific, how frequently does he or she publish sole author papers? Or sole or first author papers? And, in cases in which the last author position is honorary (and this is not universal by any means), how frequently does this author appear as sole, first or last author?

In that sense, high values of P or C (for each of their 3 variants) for an author may mean that being prolific or having many paper citations may mainly be due to the initiative of that author. By contrast, and even if an author's first order statistics are high, low values of P or C (again for each of their 3 variants) may mean that being prolific or having many paper citations may mainly be due to the initiative of that scientist's co-authors.

As regards question (c), we note that the values of C and P, as defined above, are not necessarily the same, because they are drawn from different populations. As an example from the database, a highly cited author was found to be sole, first or last author in 54% *of all his papers*. But *across his citations*, only 42% represent papers in which he was the sole, first or last author. So for these papers this author was cited less frequently than he could be found as sole, first or last author across the population of his papers. In another example, another highly cited author was found sole author in 25% of his papers. But across his citations, papers in which he was sole author represented 53% of his citations. So for these papers this author was cited more frequently than he could be found as sole author in the population of his papers.

The values of C and P being of interest in and of themselves, for a specific scientist we can also define what we call the *"claim to fame ratio"* (for lack of a better name) as the ratio $R = C/P$ with C and P as defined above (for each of their 3 variants). R can be considered as a *third-order* citations statistic, being the ratio of two second-order statistics. A value of R higher than 1 may mean that this author's prominence (or fame) is more visible within the spectrum of his or her citations than within the range of his or her papers. The opposite is the case with a low value of R, and especially if R is much lower than 1. Of course, there may very well be valid reasons for R to be much lower than 1 on a systematic basis, and calculating this statistic does not necessarily shed light into these reasons.

Note that R_s is undefined if an author has zero sole-author papers, as in this case both the C_s and P_s percentages are zero. To my surprise, there are some 895 authors in the 100,000+ scientists database *who have never authored a paper by themselves.*

Drawing from the 100,000+ scientist database, in the Appendix (Tables 1 and 2) I have calculated the 3 variants of the C, P, and R statistics for those scientists whose *primary subfields* are either (a) Logistics & Transportation, or (b) Operations Research. Citations statistics exclude self-citations. I confess that the main reason I chose these statistics to appear in this paper is that Logistics & Transportation happens to be my own primary subfield (among the 176 listed), even though they can be calculated for the entire database. A secondary reason is that Operations Research happens to be my *secondary* subfield (even though, as a result, I do not figure out in the list of people whose *primary* subfield is Operations Research). Note that the primary and secondary subfield designations (as well as the higher-level field designations) are determined by the database and *are only obliquely* the result of a scientist's own choice. This is so as these designations are defined by the predominance of journals that a given scientist has published in, which are of course his or her own choice. My own higher-level field turns out to be Economics & Business, which has a 99-percentile citation of 3,719 and is ranked as No. 11 among the 22 higher-level fields of the 100,000+ scientists database if one is to rank them by 99-percentile citations. It encompasses some 2,073 scientists in the database (among 108,277 in the 6,880,389 population). As such, and according to this criterion, in the database it ranks lower than fields such as Physics & Astronomy, Biology, and Chemistry, among others, and higher than Engineering, Mathematics & Statistics, and Philosophy & Theology, among others.

As regards the Logistics & Transportation subfield, and according to the database, this subfield has a 99-percentile citation of 1,997, listed as No. 139 among the 176 subfields in the database if one is to rank them by 99-percentile citations. As such, and according to this criterion, in the database it ranks lower than subfields such as Entomology, Ornithology, Nursing, Veterinary Sciences, and Applied Ethics, among others. It ranks higher however than subfields such as Civil Engineering, Aeronautics and Astronautics, and General Mathematics. The 99-percentile citation of 1,997 means it is futile to compare myself with a nuclear or particle physicist, whose 99-percentile citation is about 16 times higher, and it is probably also futile to compare myself to an ornithologist, a veterinarian, or an applied ethics scientist.

It turns out that 15,386 people (out of the 6,880,389 population) have Logistics & Transportation as their primary subfield, and that 84 of these people are in the top 100,000+ list. Table 1 (see Appendix) presents the top 50 scientists on this list, together with their three variants of C, P and R values (as per above), and their secondary subfield. Interestingly enough, *there are only 16 other people* in the entire 100,000+ scientists database (of which 6 in Table 1) who have both primary and secondary subfields the same as mine. This attests to the wide diversity of the database and to the difficulty of comparing scientists with one another. The list of secondary subfields for scientists who have Logistics & Transportation as their primary subfield is also very diverse, and includes Economics, Geography, Urban & Regional Planning, Human Factors, Public Health, and even Fluids and Plasmas (see Table 1).

A similar analysis can be made for the set of scientists who have Operations Research as their primary subfield. The 100,000+ database lists 319 such scientists, drawn from a group of 20,758 scientists among the 6,880,389 scientists wider group. As a primary subfield, Operations Research has a 99-percentile citation of 3,435, being ranked No. 99 among the 176 scientific subfields, that is, higher than Logistics & Transportation. It ranks however lower, according to this criterion, than subfields such as Economics, Gerontology, Sports Sciences, Optics, and Dentistry, among several others.

Table 2 (see again Appendix) lists the top 50 scientists with Operations Research as their primary subfield, together with their three variants of C, P and R values (as per above) and their secondary subfield. Interestingly enough, some of these people have Logistics & Transportation as their secondary subfield, among many others.

Both Tables 1 and 2 confirm the wide diversity in citations statistics, even across classes of scientists who have what may look like very similar expertises (incidentally, I can name at least one person in either table who has passed away). I can find no discernible pattern in either table, particularly as regards a possible correlation between the C, P and R values vis-à-vis the rank of the scientists in the database. Moreover, I find the diversity in these numbers, and especially as regards the sole author paper percentages C_s and P_s, as remarkable.

I have spent some time digging up and calculating statistics of this sort. I am sure that more can be produced, even though the value of such analyses (if any) is not immediately clear. Irrespective of this, I can only be amazed by some of the numbers. For instance, I found most intriguing that a researcher at the very top tier of the 100,000+ people database published something like 142 papers in 2020 (source: Google Scholar). This amounts to publishing around 12 new papers per month, or a new paper every 2.6 calendar days. I will refer to this researcher –who is very real- by the fictitious nickname of Galore. The question that immediately came to mind is, how much time (days, hours, minutes) did Galore spend per paper? Whenever I write a paper, even if I am not the main author, I do spend some time, so I will never reach these numbers, and I can not even remotely approach them.

So I tried to make what turned out to be an impossible calculation: estimate how many hours Galore may have spent *per paper*. Below is an attempt, which obviously involves several subjective assumptions that can be contested. Changing these assumptions will certainly change the results, however I suspect that the gist of the conclusions will not change much.

The total number of hours in a year is 8760 (8784 in 2020 as it was a leap year). Of course, nobody can work 24/7. One has to sleep, eat, go on vacation, go fishing on weekends, and engage in other non-professional activities.

The total *working hours* in a year are assumed to be 1550. At least this is so in Denmark, where I now work. I suspect it should not be very different elsewhere. Of course, no scientist dedicated to his or her work is bound by a strict working schedule, or punches a time card to work, but we need to assume a rough number to delineate between work and non-work.

Then I have assumed that the percentage of Galore's working time that was spent working on papers is 50% (the other 50% being teaching, seminars, writing proposals, going to conferences, supervising students, being engaged in administrative activities, giving interviews on TV, reviewing papers of others, etc). So according to the above, Galore spent 775 hours in 2020 working on papers that ended up being published. In my opinion, 775 hours per year likely over-represents the average scientist's time that is spent on papers.

There is more. I also had to make an assumption on how many *additional*, other than these 142 papers that were eventually published, Galore worked on during 2020. Unless we assume that Galore and his or her co-authors had a perfect publication success ratio (100%,) there may have been some additional papers that were submitted, rejected, and never resubmitted. Arbitrarily assuming a success ratio of 80%, this means that Galore worked on 142/0.8 = 177 papers in 2020. Dividing 775 hours by 177, one yields an average of 4.38 hours (or 4 hours and 23 minutes) of work per paper. Assuming a lower success ratio will result in a lower time per paper.

Needless to say, these 4 hours and 23 minutes per paper also include, in addition to paper writing, editing, revision and other paper-writing tasks, they include all the actual prior research that led to each of these papers, including data collection, model formulation, computational runs, lab experiments, field surveys, validation, etc. This assumes of course that Galore contributed at least to some the above activities, and even though some of them may have taken place in prior years.

As 4 hours and 23 minutes per paper looks incredibly low, I next looked at the database to compute Galore's C, P and R statistics, and obtained the following approximate picture:

C_s=9%, C_{sf}=20%, C_{sfl}=38%
P_s=13%, P_{sf}=28%, P_{sfl}=64%
R_s=0.67, R_{sf}=0.72, R_{sfl}=0.60

Most C and P percentages being low and all R "claim-to-fame" ratios being less than 1, this gives the clear impression that much (but surely not all) of Galore's bibliometric performance can be attributed to the work of his or her co-authors. This may also explain the 4 hours and change. Of course, being able to assemble a broad network of scientific collaborators is very important, in and of itself. And for sure these results cannot be generalized to other scientists. However, I wonder how any minimal level of quality control can be maintained, either at the research level, or at the paper production level, or finally at the paper review level, if these numbers are valid.

4. Conclusions

To conclude, and even though bibliometrics is an interesting sport, and if one wants to look below the surface some interesting information may be revealed, in my view the importance of bibliometrics is way far overblown. I may want to write a paper that I want to keep for myself, or share with only a few people, or never bother publishing it the traditional way. What I think is more important for any given paper is whether its content is sound, whether it has improved upon the state of the art, whether it has developed a new method, whether its results are useful to science, industry or society, whether it has provided insights, and in general whether it has stimulated more research in an area. In addition, some parts of an academic's professional life, such as for instance spending time with industry, may not be reflected at all in any citation metric, even though they may be just as important. This paper has also shown the difficulties in comparing scientists with one another.

Therefore, and as far as I am concerned, we can live without bibliometrics.

REFERENCES

Aad, G. et al. 2015 (ATLAS Collaboration, CMS Collaboration) *Physical Review Letters* 114, 191803.

Archambault, E., Caruso, J., & Beauchesne, O. , 2011. Towards a multilingual, comprehensive and open scientific journal ontology. Proceedings of the 13th International Conference of the International Society for Scientometrics and Informetrics, 66–77.

Barabasi, A-L., Albert. R., 1999, The emergence of scaling in random networks, *Science* , Vol. 286, Issue 5439, 509-512 DOI: 10.1126/science.286.5439.509

Chang, Y.-T., Choi. K.-S., Jo. A., Park. H., 2018. Top 50 authors, affiliations and countries in maritime research. *Int. J. Shipping and Transport Logistics*. Vol. 10. No. 1. 87-111.

Golden. B.L., Shier. D.R., 2020. Twenty-one years in the life of *Networks* (2000 to 2020). *Networks*. 2020;1–9. DOI: 10.1002/net.21986

Ioannidis. J.P.A., 2005. Why Most Published Research Findings Are False. *PLoS Medicine*. https://doi.org/10.1371/journal.pmed.0020124

Ioannidis JPA, Baas J. Klavans R. Boyack KW (2019) A standardized citation metrics author database annotated for scientific field. *PLoS Biology* 17(8): e3000384. https://doi.org/10.1371/journal. pbio.3000384

Ioannidis. J.P.A. 2020. Global perspective of COVID-19 epidemiology for a full-cycle pandemic. *European Journal of Clinical Investigation*. https://doi.org/10.1111/eci.13423

Jiang, C., Bhat, C., and Lam, W., 2020. A bibliometric overview of *Transportation Research Part B: Methodological* in the past forty years (1979–2019). *Transportation Research Part B: Methodological* 138, 268–291.

Psaraftis, H.N., 1984. The practical importance of asymptotic optimality in certain heuristic algorithms. *Networks* 14. No.4. 587-596.

Sahoo, S. and Schönborn. A., 2020. A bibliometric overview of WMU Journal of Maritime Affairs since its inception in 2002. *WMU Journal of Maritime Affairs* https://doi.org/10.1007/s13437-020-00197-w

APPENDIX

Table 1: Top 50 scientists who have Logistics & Transportation as their primary subfield, together with their three C, P and R variants and their secondary subfield. Source: own analysis based on Ioannidis et al. (2019).

No.	Scientist	Rank in database	C_s (%)	P_s (%)	R_s	C_{sf} (%)	P_{sf} (%)	R_{fs}	C_{sfl} (%)	P_{sfl} (%)	R_{sfl}	Secondary subfield
1	Daganzo, Carlos F.	2,689	56.00	36.14	1.549	69.66	53.01	1.314	97.87	92.17	1.062	Operations Research
2	Hensher, David A.	3,391	13.68	20.78	0.659	57.61	53.88	1.069	87.47	84.25	1.038	Economics
3	Cervero, Robert	4,391	23.47	56.03	0.419	64.86	84.40	0.768	85.11	90.78	0.938	Urban & Regional Planning
4	Bhat, Chandra R.	6,078	26.65	13.75	1.938	53.46	36.67	1.458	83.05	72.08	1.152	Geography
5	Flyvbjerg, Bent	6,369	56.98	42.31	1.347	90.65	71.79	1.263	92.78	82.05	1.131	Urban & Regional Planning
6	Train, Kenneth	6,564	52.77	24.64	2.142	58.43	52.17	1.120	95.66	91.30	1.048	Energy
7	Yang, Hai	9,212	5.96	3.21	1.859	49.33	25.00	1.973	74.73	63.46	1.178	Operations Research
8	Handy, Susan	11,556	12.00	12.60	0.953	45.34	29.92	1.515	81.87	76.38	1.072	Urban & Regional Planning
9	Mokhtarian, Patricia L.	12,712	8.99	10.24	0.878	30.91	27.11	1.140	75.57	74.10	1.020	Urban & Regional Planning
10	Banister, David	16,770	29.16	30.72	0.949	46.21	46.99	0.983	68.34	83.13	0.822	Urban & Regional Planning
11	Ewing, Reid	16,929	12.11	36.09	0.336	69.70	65.09	1.071	75.39	85.21	0.885	Urban & Regional Planning
12	Williams, Allan F.	17,807	13.88	13.18	1.053	40.51	50.00	0.810	75.03	76.36	0.983	Public Health
13	Bell, Michael G.H.	19,015	21.62	14.85	1.456	40.29	28.22	1.428	82.48	66.83	1.234	Numerical & Computational Mathematics
14	Rietveld, Piet	19,482	4.56	9.31	0.490	12.87	17.72	0.726	79.30	79.28	1.000	Economics
15	Papageorgiou, Markos	20,557	7.08	8.70	0.815	35.47	20.07	1.767	64.40	64.55	0.998	Industrial Engineering & Automation
16	Elvik, Rune	23,067	76.48	71.32	1.072	88.64	82.35	1.076	97.93	94.12	1.041	Human Factors
17	Abdel-Aty, Mohamed	24,607	4.93	2.36	2.086	41.62	27.03	1.540	79.88	61.15	1.306	Human Factors
18	Mannering, Fred L.	26,805	3.00	6.25	0.481	15.67	21.53	0.728	85.83	82.64	1.039	Civil Engineering
19	Levinson, David	33,196	25.97	14.59	1.779	42.61	30.81	1.383	86.82	92.43	0.939	Urban & Regional Planning
20	van Wee, Bert	33,776	8.66	10.40	0.833	18.31	25.74	0.711	78.51	81.68	0.961	Urban & Regional Planning
21	Mahmassani, Hani S.	35,371	1.94	1.64	1.181	21.73	15.74	1.381	76.08	62.62	1.215	Operations Research
22	Notteboom, Theo	35,534	25.08	19.84	1.264	64.72	43.65	1.483	93.77	80.16	1.170	Geography
23	Sheu, Jiuh Biing	36,011	53.13	36.08	1.473	79.14	62.89	1.258	90.15	80.41	1.121	Operations Research

24	Pucher, John	36,211	6.57	33.82	0.194	82.10	77.94	1.053	90.19	89.71	1.005	Urban & Regional Planning
25	Evans, Leonard	37,374	61.33	57.66	1.064	91.47	87.39	1.047	96.44	94.59	1.019	Human Factors
26	Viano, David C.	37,396	6.03	12.24	0.492	38.37	48.64	0.789	52.18	71.43	0.730	Neurology & Neurosurgery
27	Timmermans, Harry J.P.	38,336	2.32	3.30	0.704	10.25	8.58	1.194	85.54	83.33	1.026	Geography
28	Noland, Robert B.	39,947	9.64	10.20	0.945	34.99	35.37	0.989	76.20	68.03	1.120	Energy
29	Golob, Thomas F.	40,344	21.97	8.45	2.599	66.18	60.56	1.093	85.42	84.51	1.011	Geography
30	Zhang, H. Michael	41,740	28.26	10.27	2.750	43.26	22.60	1.914	82.72	78.77	1.050	General Science & Technology
31	Summala, Heikki	42,596	10.27	11.11	0.925	24.23	24.24	1.000	89.00	83.84	1.062	Human Factors
32	Kockelman, Kara M.	43,320	6.74	5.96	1.130	15.30	13.91	1.100	82.61	70.86	1.166	Urban & Regional Planning
33	Lo, Hong K.	44,538	7.63	2.62	2.916	43.06	25.13	1.714	81.85	76.96	1.063	Operations Research
34	Verhoef, Erik T.	45,217	16.84	14.29	1.179	40.71	32.54	1.251	88.57	80.16	1.105	Economics
35	Shinar, David	45,813	10.97	13.21	0.831	50.81	43.40	1.171	83.67	80.19	1.043	Human Factors
36	Wong, Sze Chun	46,880	2.76	2.48	1.112	14.96	13.66	1.095	53.34	53.11	1.004	Applied Mathematics
37	Ben-Akiva, Moshe	48,874	0.25	1.29	0.190	31.73	23.28	1.363	78.77	79.31	0.993	Marketing
38	Airey, Gordon D.	53,301	31.49	6.06	5.196	57.22	20.61	2.777	74.01	53.94	1.372	Building & Construction
39	Psaraftis, Harilaos N.	53,647	42.46	25.41	1.671	63.80	46.72	1.366	82.95	83.61	0.992	Operations Research
40	Axhausen, Kay W.	53,857	3.42	4.78	0.714	14.95	12.92	1.157	68.11	76.08	0.895	Urban & Regional Planning
41	Masad, Eyad	55,018	1.22	0.38	3.205	35.46	20.23	1.753	51.92	43.51	1.193	Building & Construction
42	Hoogendoorn, Serge P.	57,979	0.96	0.87	1.106	36.57	15.32	2.387	73.57	72.25	1.018	Fluids & Plasmas
43	Nagel, Kai	58,002	8.82	5.13	1.721	40.47	17.95	2.255	70.67	84.62	0.835	Fluids & Plasmas
44	Lord, Dominique	58,415	5.34	4.20	1.271	47.90	22.69	2.111	77.90	68.07	1.145	Human Factors
45	Huang, Hai Jun	59,806	2.37	1.26	1.875	12.95	9.46	1.368	52.83	52.37	1.009	General Physics
46	Karlaftis, Matthew G.	60,094	5.70	5.95	0.958	30.66	26.79	1.145	66.74	64.88	1.029	Civil Engineering
47	Hall, Randolph W.	60,575	47.29	52.69	0.898	60.26	76.34	0.789	73.17	90.32	0.810	Operations Research
48	Shope, Jean T.	60,791	8.87	9.26	0.958	30.27	28.70	1.055	52.91	51.85	1.020	Public Health
49	Quddus, Mohammed A.	61,882	8.03	5.32	1.510	44.16	22.34	1.977	69.00	43.62	1.582	Human Factors
50	Hauer, Ezra	63,248	49.49	54.29	0.912	93.38	84.29	1.108	96.45	90.00	1.072	General Arts, Humanities & Social Sciences

Table 2: Top 50 scientists who have Operations Research as their primary subfield, together with their three C, P, and R variants and their secondary subfield. Source: own analysis based on Ioannidis et al. (2019).

No.	Scientist	Rank in database	C_s (%)	P_s (%)	R_s	C_{sf} (%)	P_{sf} (%)	R_{fs}	C_{sfl} (%)	P_{sfl} (%)	R_{sfl}	Secondary subfield
1	Saaty, Thomas L.	1,543	72.90	46.45	1.569	91.64	83.87	1.093	99.65	96.13	1.037	Numerical & Computational Mathematics
2	Laporte, Gilbert	1,633	7.79	3.72	2.095	21.19	15.91	1.332	63.41	63.02	1.006	Logistics & Transportation
3	Sarkis, Joseph	2,043	12.40	11.73	1.057	28.82	29.91	0.963	66.41	70.67	0.940	Business & Management
4	Gunasekaran, Angappa	4,041	7.00	9.96	0.703	50.10	34.15	1.467	67.29	58.33	1.154	Business & Management
5	Glover, Fred	4,050	37.73	11.40	3.310	60.09	43.32	1.387	80.05	72.64	1.102	Artificial Intelligence & Image Processing
6	Cheng, T.C.E.	4,132	4.59	9.78	0.469	30.96	31.51	0.982	73.68	74.39	0.991	Computation Theory & Mathematics
7	Banker, Rajiv D.	4,138	7.09	4.35	1.630	94.88	80.12	1.184	95.34	86.34	1.104	Accounting
8	Lee, Hau L.	4,740	14.41	8.93	1.614	70.49	47.32	1.489	92.13	81.25	1.134	Logistics & Transportation
9	Kusiak, Andrew	4,907	12.85	19.64	0.654	62.87	55.65	1.130	94.53	90.18	1.048	Energy
10	Bertsekas, Dimitri P.	5,881	26.92	34.67	0.776	53.00	60.30	0.879	85.26	91.46	0.932	Industrial Engineering & Automation
11	Goyal, Suresh	6,518	31.21	35.59	0.877	59.17	54.92	1.077	88.76	86.78	1.023	Business & Management
12	Mangasarian, Olvi L.	6,645	19.83	36.42	0.544	48.56	63.58	0.764	95.77	94.04	1.018	Artificial Intelligence & Image Processing
13	Whitt, Ward	6,973	37.29	30.86	1.209	37.41	31.97	1.170	99.82	98.51	1.013	Statistics & Probability
14	Chan, Felix T.S.	7,002	7.62	5.88	1.295	57.91	40.95	1.414	81.02	63.80	1.270	Industrial Engineering & Automation
15	Beasley, J. E.	10,269	34.67	24.21	1.432	56.68	40.00	1.417	91.71	90.53	1.013	Bioinformatics
16	Van Wassenhove, Luk	10,627	3.50	0.45	7.698	5.80	4.55	1.277	89.91	83.18	1.081	Business & Management
17	Mingers, John	10,638	60.01	50.79	1.181	91.19	76.19	1.197	98.76	93.65	1.055	Information Systems
18	Cachon, Gérard P.	10,996	26.71	26.19	1.020	92.53	88.10	1.050	99.06	92.86	1.067	Marketing
19	Lee, Chung Yee	11,140	11.10	4.91	2.262	44.73	35.58	1.257	84.70	80.98	1.046	Logistics & Transportation
20	Tseng, Paul	11,492	44.35	34.29	1.294	55.86	50.48	1.107	91.91	91.43	1.005	Industrial Engineering & Automation
21	Fisher, Marshall L.	12,125	25.44	22.95	1.108	48.52	65.57	0.740	89.68	85.25	1.052	Computation Theory & Mathematics
22	Gendreau, Michel	13,168	1.30	2.50	0.522	24.86	16.88	1.473	47.77	48.44	0.986	Logistics & Transportation

23	Zhu, Joe	13,178	17.00	9.80	1.734	17.48	10.46	1.671	94.54	90.85	1.041	Economics
24	Towill, Denis R.	13,377	13.55	27.53	0.492	20.88	38.87	0.537	91.68	85.83	1.068	Business & Management
25	Ngai, Eric W.T.	13,386	3.23	2.16	1.495	48.95	33.51	1.461	81.51	75.68	1.077	Business & Management
26	Bertsimas, Dimitris	13,577	1.00	2.14	0.467	93.62	86.10	1.087	97.56	95.19	1.025	Statistics & Probability
27	Combettes, Patrick L.	14,623	24.66	18.40	1.340	75.96	67.20	1.130	88.54	83.20	1.064	General Mathematics
28	Sherali, Hanif D.	15,623	1.72	5.11	0.336	26.47	43.55	0.608	51.69	79.03	0.654	Logistics & Transportation
29	Bard, Jonathan F.	15,884	18.27	19.32	0.946	59.62	55.11	1.082	77.84	76.70	1.015	Logistics & Transportation
30	Cooper, William W.	15,943	0.28	7.38	0.038	14.69	41.61	0.353	41.99	53.69	0.782	Economics
31	Fukushima, Masao	17,304	18.36	11.64	1.577	28.89	19.58	1.475	93.87	88.89	1.056	Numerical & Computational Mathematics
32	Kleijnen, Jack P.C.	17,417	40.94	37.18	1.101	76.32	67.95	1.123	92.74	88.46	1.048	Statistics & Probability
33	Pang, Jong Shi	17,578	9.49	14.66	0.648	31.00	36.65	0.846	88.21	81.15	1.087	Networking & Telecommunications
34	Qi, Li qun	18,029	12.07	9.15	1.320	41.53	23.86	1.741	73.09	66.34	1.102	Numerical & Computational Mathematics
35	Brucker, Peter	18,283	23.06	17.12	1.347	78.58	85.59	0.918	98.27	90.09	1.091	Computation Theory & Mathematics
36	Wright, Stephen J.	18,288	8.31	18.99	0.438	20.10	25.32	0.794	85.96	81.01	1.061	Numerical & Computational Mathematics
37	Tang, Christopher S.	19,440	18.60	9.77	1.903	35.97	24.06	1.495	96.59	87.97	1.098	Business & Management
38	Nesterov, Yurii	19,568	58.44	33.64	1.737	80.14	60.75	1.319	90.78	80.37	1.129	Numerical & Computational Mathematics
39	Hochbaum, Dorit S.	20,229	19.69	24.63	0.799	85.13	73.88	1.152	94.54	86.57	1.092	Computation Theory & Mathematics
40	Hansen, Pierre	20,425	0.78	2.43	0.322	44.73	44.68	1.001	75.69	69.30	1.092	Computation Theory & Mathematics
41	Keeney, Ralph L.	20,722	45.97	43.41	1.059	73.53	75.97	0.968	88.90	92.25	0.964	Strategic, Defence & Security Studies
42	L'Ecuyer, Pierre	20,845	32.35	18.91	1.711	66.92	48.76	1.373	83.27	76.62	1.087	Numerical & Computational Mathematics
43	Taillard, Eric D.	20,995	45.81	25.00	1.832	57.40	44.44	1.292	86.20	86.11	1.001	Logistics & Transportation
44	Kao, Chiang	21,067	26.86	31.39	0.856	89.56	89.78	0.997	98.06	97.08	1.010	Artificial Intelligence & Image Processing
45	Goldfarb, Donald	21,228	19.83	11.11	1.785	44.60	63.33	0.704	60.09	82.22	0.731	Numerical & Computational Mathematics
46	Lasserre, Jean Bernard	21,246	53.59	43.62	1.228	61.47	59.26	1.037	92.07	89.30	1.031	Industrial Engineering & Automation
47	ReVelle, Charles	21,426	4.75	4.65	1.021	32.72	22.67	1.443	85.48	72.67	1.176	Environmental Engineering

48	Cordeau, Jean François	21,601	2.46	0.74	3.320	53.07	26.67	1.990	56.06	32.59	1.720	Logistics & Transportation
49	Guide, Jr. V. Daniel R.	21,985	11.38	6.76	1.685	71.25	58.11	1.226	77.07	78.38	0.983	Business & Management
50	Dekker, Rommert	22,047	9.47	2.58	3.669	27.27	18.71	1.457	54.50	65.16	0.836	Logistics & Transportation