How to avoid the ‘Jesus multiplying bread and fish’ myth and ‘Münchausen baron bootstrap’ legend in individual bibliometric indices: a simple bimedian method applied to Italian ASPA academic members

Giuseppe Pulina

To cite this article: Giuseppe Pulina (2020) How to avoid the ‘Jesus multiplying bread and fish’ myth and ‘Münchausen baron bootstrap’ legend in individual bibliometric indices: a simple bimedian method applied to Italian ASPA academic members, Italian Journal of Animal Science, 19:1, 644-649, DOI: 10.1080/1828051X.2020.1777753

To link to this article: https://doi.org/10.1080/1828051X.2020.1777753
Dear Editor in Chief,

Over a decade ago we assessed the status of the bibliographic indicators of ASPA members (Pulina and Francesconi 2007) and since then much water has passed under the bridge of bibliometric science. The number of published papers, the rank of scientific journals, number of citations and Hirsch’s H-index (Hirsch 2005), have been widely used for the evaluation of academic careers and of scientific institutions (El Gibari et al. 2018). The Italian National Agency for the Evaluation of Universities and Research Institutes (ANVUR 2020) has adopted three bibliometric indices for the ranking of the single academic careers in a given period: the total number of published papers, the total citations, and the H-index. The median of each of the three indices has been used as a threshold value. Thus, for example, for becoming a university professor, that a candidate must have at least two out of the three indicators above the threshold value.

The weakness of the abovementioned indicators to evaluate the single scientific career is well known. Limitations of the H-index, originally proposed to assess the real impact of the work of a single researcher on the scientific community, and of its variants for different scientific fields have been reviewed by Alonso et al. (2009). However, multiple authorship and self-citations still represent two main issues for the correct evaluation of the individual scientific career.

Galam (2011) proposed a gh-index to account for the multiple authorship problem and observed that:

> ‘When a measure becomes a target, it ceases to be a good measure’

Galam (2011) proposed a gh-index to account for the multiple authorship problem and observed that:

> ‘When a measure becomes a target, it ceases to be a good measure’

The co-authorship H-index inflation was faced by Hirsch (2010) himself who modified his indicator to consider for multiple authors and tailored by Prathap (2011) who proposed a fractional and harmonic p-indices for multiple authorship.

Self-citations inflate the bibliometric indices too. Recently, SCOPUS® Database has beginning to release two H-index corrected for the self-citation by the specific author and of all its co-authors. This theme has been developed at country level by Baccini et al. (2019) who, paraphrasing Goodhart’s law (‘when a measure becomes a target, it ceases to be a good measure’) observed that ‘the introduction of bibliometric targets into the Italian academy has changed the way Italian academics cite and use the work of their colleagues’.

To overcome the problems of multiple co-authorship and self-citations and to avoid the use of composite indicators difficult to calculate and validate (OECD 2008), I propose here a simple bi-median method for the evaluation of a single career of the researchers. The method is based on the indices currently available in the most popular bibliometric databases (SCOPUS® and WOS®) and that are widely used by the Italian and international scientific communities. I then have applied this method to Italian academic ASPA members.

Data were retrieved from the net during the period from 20th April to 30th April 2020. The list of names and academic position of 365 ASPA members (Full professor, Associate professor, Senior researcher, and Junior researcher; the difference between researcher position is that the former has a permanent tenure and the latter temporary tenure) were downloaded from the website of the ‘Ministero dell’Istruzione, dell’Università e della Ricerca’ (the Italian Ministry of Education and Research, MIUR) [https://cercauniversita.cineca.it/php5/docenti/cerca.php]. Members were grouped according to the four research areas of ASPA (Scientific Disciplinary Sectors, S.D.S.): (i) AGR/17 – Animal genetics and breeding (’Zootechnica Generale e Miglioramento Genetico’); (ii) AGR/18 – Animal nutrition and feeding (’Nutrizione e Alimentazione Animale’); (iii) AGR/19 – Animal husbandry (’Zootechnica Speciale’); (iv) and AGR/20 – Poultry, rabbits and fish production (’Zoocolture’).

Then the publication list of all academic members (17 data were missed) and their bibliometric indices were...
downloaded from the SCOPUS® Database [https://www.elsevier.com/solutions/scopus]. Considered indices were: total number of publications (Papers n.), total citations (Cit. tot (1)), total citations without self-citations of all co-authors (Cit. no self_cit_all (2)), total citation without self-citations of the considered author (Cit. no self_cit_aut (3)), H(1) index calculated over Cit. tot (1), H(2) index calculated excluding self-citation of all co-authors, H(3) index calculated excluding self-citations of the selected author, total number of co-authors (Co-aut tot.). Then, I calculated the following ratio indices: co-authors per paper (Co-aut paper ¼ Co-aut tot./Papers n.), citation per paper (Cit/paper ¼ Cit. tot (1)/Papers n.), citation per author (Cit/aut_all ¼ Cit. tot (1)/Co-aut tot.), incidence of no self-citations of all co-authors over total citations (% self_cit_all (2) ¼ Cit. no self_cit_all (2)/Cit. tot (1) × 100), incidence of no self-citations of selected author over total citations (% self_cit_aut (3) ¼ Cit. no self_cit_aut (3)/Cit. tot (1) × 100) and the incidence of H(2) and H(3) over H(1).

Table 1 reports the median, the first (Q1) and third (Q3) quartile of each index and index ratio calculated for the Italian four academic positions and the 348 ASPA academic members.

Comparing the data shown in Table 1 with those we reported 14 years ago (Pulina and Francesconi 2007), the overwhelming increase in all values is undeniable: Papers n., 7.6 times; Cit. tot (1), 32.7 times; H(1), 6.5 times; Cit/paper, 4.5 times. This confirms the commitment of the ASPA scientific community to improve its position in the international panorama of animal science research.

Table 1. Medians, the first (Q1) and third (Q3) quartiles of bibliometric indices and index ratios calculated for a single academic position and for all ASPA academic members.

| Variable          | Measure                  | Full professor  (n = 85) | Associate professor (n = 140) | Senior researcher (n = 76) | Junior researcher (n = 47) | All (n = 348) |
|-------------------|--------------------------|--------------------------|------------------------------|---------------------------|---------------------------|---------------|
| Papers n          | Q1                       | 52                       | 35                           | 17                        | 25                        | 28            |
|                   | Median                   | 70                       | 48                           | 26                        | 30                        | 46            |
|                   | Q3                       | 95                       | 66                           | 37                        | 42                        | 64            |
| Cit. tot (1)      | Q1                       | 696                      | 357                          | 155                       | 154                       | 261           |
|                   | Median                   | 1190                     | 586                          | 263                       | 246                       | 552           |
|                   | Q3                       | 1721                     | 945                          | 448                       | 541                       | 1016          |
| Cit. no self_cit_all (2) | Q1                       | 529                      | 252                          | 122                       | 87                        | 189           |
|                   | Median                   | 925                      | 434                          | 207                       | 152                       | 419           |
| Cit. no self_cit_aut (3) | Q1                       | 615                      | 307                          | 142                       | 113                       | 230           |
|                   | Median                   | 1044                     | 509                          | 246                       | 186                       | 490           |
| H (1)             | Q1                       | 15                       | 10                           | 6                         | 7                         | 9             |
|                   | Median                   | 19                        | 13                           | 9                         | 9                         | 13            |
|                   | Q3                       | 23                       | 17                           | 12                        | 13                        | 18            |
| H (2)             | Q1                       | 13                       | 9                            | 6                         | 5                         | 7             |
|                   | Median                   | 17                        | 11                           | 7                         | 7                         | 11            |
| H (3)             | Q1                       | 20                       | 15                           | 10                        | 10                        | 15            |
|                   | Median                   | 18                        | 12                           | 9                         | 8                         | 12            |
| Co-aut tot.       | Q1                       | 97                       | 75                           | 41                        | 52                        | 62            |
|                   | Median                   | 136                      | 106                          | 61                        | 76                        | 98            |
| Co-aut paper      | Q1                       | 1.56                     | 1.70                         | 1.86                      | 1.92                      | 1.70          |
|                   | Median                   | 1.90                      | 2.13                         | 2.55                      | 2.46                      | 2.18          |
| Cit/paper         | Q1                       | 12.22                    | 8.75                         | 7.60                      | 5.79                      | 8.47          |
|                   | Median                   | 15.19                     | 12.30                        | 9.44                      | 7.73                      | 12.17         |
| Cit/aut_all       | Q1                       | 5.67                     | 3.64                         | 2.52                      | 2.02                      | 3.20          |
|                   | Median                   | 7.71                      | 5.48                         | 3.73                      | 2.97                      | 5.48          |
| % self_cit_all (2) | Q1                       | 3.91                     | 9.08                         | 6.58                      | 5.67                      | 8.77          |
|                   | Median                   | 11.01                     | 15.98                        | 14.97                     | 12.42                     | 16.49         |
| % self_cit_aut (3) | Q1                       | 85.45                    | 81.70                        | 86.05                     | 75.10                     | 81.80         |
|                   | Median                   | 89.84                     | 89.39                        | 93.92                     | 84.49                     | 89.59         |
| H(2)/H(1)         | Q1                       | 93.12                    | 94.14                        | 96.06                     | 88.89                     | 94.07         |
|                   | Median                   | 89.84                     | 89.39                        | 93.92                     | 84.49                     | 89.59         |
| H(3)/H(1)         | Q1                       | 0.91                     | 0.87                         | 0.90                      | 0.80                      | 0.87          |
|                   | Median                   | 0.94                      | 0.93                         | 1.00                      | 0.89                      | 0.93          |

Data were collected on SCOPUS® platform from April 20th to April 30th, 2020.
Figure 1 shows the position of ASPA academic members in the space defined by the number of papers and total citations (left) and $H(1)$ (right). The medians of indicators (red lines) define 4 quadrants: low-low (LL, left-bottom), high-low (HL, right-bottom), high-high (HH, right-upper) and low-high (LH, left-upper). Being the indices highly correlated, as it can be seen next in this letter, almost all members are located in the LL and HH quadrants. The few observations falling in the LH quadrant refer belong to scientists with limited published papers having a high impact, whereas those in HL quadrants belong to scientists with many published papers but with scarce citations, respectively. No significant differences were found among different disciplinary sectors ($p > .05$, data not shown), thus confirming that ASPA appears as a quite homogeneous scientific community regardless of the specific field of interest of each member.

The correlation matrix of the main bibliometric indicators and their derived indices is reported in Table 2. Papers n. shows a low negative correlation with Co-aut paper and %self_cit_aut (3), whereas it is positively, mainly strongly, correlated with the other parameters. This result suggests that the number of published papers represents the main driver of the author’s impact on the scientific community. Moreover, they suggest that these parameters should not be treated as independent when building a bimedian index, except for those that resulted in weakly correlated. The other main bibliometric parameter, that is, Cit. tot (1), and also the Cit. no self_cit_aut (3), are not correlated with Co-aut paper and %self_cit_aut (3). Based on these results, total citations with or without author self-citations can be used as an independent variable against co-authorships load and self-citations incidence, the two parameters of interest of this letter.

Based on the weak or null correlation amongst the indicators above discussed, I have developed some bimedian graphs that can help to discriminate high shared authorship and high self-citation ratio as the worse position in the bibliometric space of the scientific careers of academic ASPA members. All graphs highlighted the expected differences among the four academic positions, as the result of the different lengths of the scientific careers.

The best position of the left graph of Figure 2 is HL followed, in decreasing order by LL, HH, and LH position. A low number of co-authors means a larger quota of the intellectual property of the paper for the researcher. However, the weakness of this parameter is represented by the penalisation of the open authorship, that is, the more the co-authors the larger the network of a single researcher. I am aware that for assessing the correct amount of intellectual property of a multiple-author

![Cit tot (1) vs Papers](image1.png)

![H(1) vs Papers](image2.png)

Figure 1. ASPA scientific community divided by the scientific disciplinary sector (S.D.S.). Scatterplot of the number of papers published by each member plotted against total citations (Cit. tot (1), on the left) and $H(1)$ (on the right) indices and the relative position to medians (red scattered lines).

Table 2. Pearson correlations between some of the bibliometric indices and index ratios calculated for all ASPA academic members.

|                      | Papers n. | Cit. tot (1) | Cit. no self_cit_aut (3) | H (1) | H (3) | Co-aut paper | Cit/aut_all | % self_cit_aut (3) | % self_cit_aut (3)/Cit/aut_all |
|----------------------|-----------|--------------|--------------------------|-------|-------|--------------|-------------|---------------------|---------------------------------|
| Cit. tot (1)         | 0.836     |              |                          |       |       |              |             |                     |                                 |
| Cit. no self_cit_aut (3) | 0.778     | 0.988        |                          |       |       |              |             |                     |                                 |
| $H(1)$               | 0.861     | 0.928        | 0.911                    |       |       |              |             |                     |                                 |
| $H(3)$               | 0.812     | 0.922        | 0.983                    |       |       |              |             |                     |                                 |
| Co-aut paper         | –0.221    | –0.079       | –0.040                   | –0.162| –0.132|              |             |                     |                                 |
| Cit/aut_all          | 0.408     | 0.631        | 0.631                    | 0.640 | 0.647 | –0.489       |             |                     |                                 |
| % self_cit_aut (3)   | –0.316    | –0.095       | 0.006                    | –0.135| –0.025| 0.270        | 0.019       |                      |                                 |

Data collected on SCOPUS® platform from April 20th to April 30th, 2020. Bold and italic values indicate not significantly different from zero ($p > .05$).
paper it is necessary to consider many other aspects like the position and role of each author (first or last, corresponding, etc.), the rank of the scientific journal, or the kind of paper (review, research contribution, technical note, short paper, etc.). Thus, for the sake of simplicity, I chose to consider the average number of co-authors per paper as a proxy for the evaluation of the intellectual property of a scientist published literature.

The best position of the right graph of Figure 2 is HH, followed equally by HL and LH positions and, again worse, the LL position. Senior researchers show significantly less self-citation index compared to Full and Associate professors, that exhibit less self-citations than Junior researchers ($p < .05$, data not shown).

The left graph in Figure 3 reports the average authorship plotted against total citations. The best position is in HL quadrant, followed in order by HH, LL and LH quadrants.

Figure 2. ASPA scientific community divided by academic role. Scatterplot of the number of papers published by each member plotted against mean author per paper (co_aut_paper, on the left) and % of total citations exclude self-citations (%self_cit_aut(3), on the right) indices and the relative position to medians (red scattered lines).

Figure 3. ASPA scientific community divided by academic role. Scatterplot of the number of total citations (Cit tot (1)) by each member plotted against mean author per paper (co_aut_paper, on the left) and % of total citations excluded self-citations (%self_cit_aut(3), on the right) indices and the relative position to medians (red scattered lines).

The right graph of the same figure plots the percentage of no self-citations (%self_cit_aut(3)) over total citations. It reflects the pattern of the companion graph and therefore has the same qualitative order of quadrants.

The best position of the left graph in Figure 4 is the HH quadrant, that is, researchers with a high degree of not self-citation and high H-index calculated without self-citations followed by LH, HL and LL quadrants.

The right graph in the same figure plots the average citation per author per paper against the self-citations index. It can be interpreted with the same order of quadrants of the companion graph.

Finally, Figures 5 shows the distribution of the H-index excluded the self-citations by the researcher plotted against the average citation per author per paper for full and associate professors respectively. The best position is that of quadrant HH (high citation per co-author coupled...
Conclusions

The ASPA community exhibits a good scientific performance with high values of the main bibliometric indicators, obviously greater for the older compared to the younger scientists, with a limited number of co-authors per paper and a low degree of self-citations. However, the bimedian method I proposed here, that can be also used with other pairs of medians than those reported in this letter and enriched with Q1 or Q3 thresholds to highlight areas of excellence, could be a useful tool to evaluate our ASPA academic members careers in the light of avoiding the ‘Jesus multiplying bread and fish’ myth and ‘Münchhausen baron bootstrap’ legend.

Disclosure statement

No potential conflict of interest was reported by the author.

References

Alonso S, Cabreroiz FJ, Herrera-Viedma E, Herrera F. 2009. h-Index: a review focused in its variants, computation and standardization for different scientific fields. J Informetr. 3(4):273–289.
ANVUR. 2020. Typology of ANVUR’s activities. ANVUR website, Rome. Accessed date 10th May 2020. https://www.anvur.it/wp-content/uploads/2020/01/Typology-of-ANVUR%E2%80%99s-activities.pdf
Baccini A, De Nicolao G, Petrovich E. 2019. Citation gaming induced by bibliometric evaluation: a country-level comparative analysis. PLoS One. 14(9):e0221212.
El Gibari S, Gómez T, Ruiz F. 2018. Evaluating university performance using reference point based composite indicators. J Informetr. 12(4):1235–1250.
Galam M. 2011. Tailor based allocations for multiple authorship: a fractional gh-index. Scientometrics. 89(1):365–379.
Hirsch JE. 2005. An index to quantify an individual’s scientific research output. Proc Natl Acad Sci USA. 102(46):16569–16572.
Hirsch JE. 2010. An index to quantify an individual’s scientific research output that takes into account the effect of multiple coauthorship. Scientometrics. 85(3):741–754.
OECD. 2008. Handbook on constructing composite indicators: methodology and user guide. Paris (France): OECD publications; p. 1–158.
Prathap G. 2011. The fractional and harmonic p-indices for multiple authorship. Scientometrics, 86:239–244
Pulina G, Francesconi HED. 2007. Some bibliometric indexes for members of the Scientific Association of Animal Production (ASPA). Ital J Anim Sci. 6(1):83–103.

Giuseppe Pulina
Department of Agraria, Honorary president of Animal Science and Production Association (ASPA)

University of Sassari, Sassari, Italy

gpulina@uniss.it

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.