

Supporting information

Perovskite Device Efficiency is a Poor Predictor for the Number of Citations a Paper Will Get

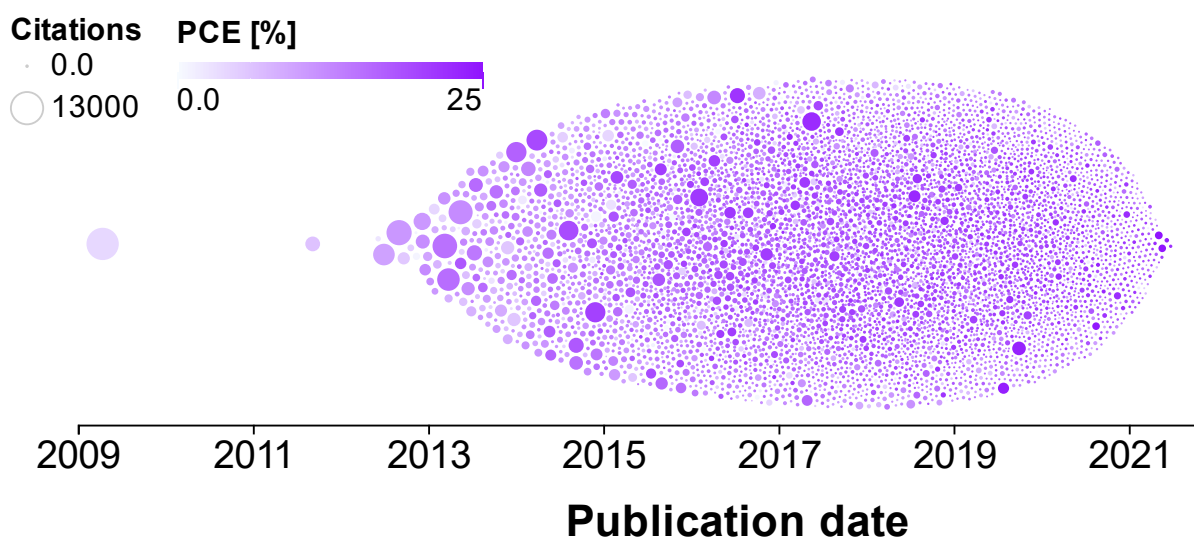
Ayman Maqsood¹, T. Jesper Jacobsson*¹

1. Institute of Photoelectronic Thin Film Devices and Technology, Key Laboratory of Photoelectronic Thin Film Devices and Technology of Tianjin, College of Electronic Information and Optical Engineering, Nankai University, Tianjin 300350, China

Corresponding author:

T. Jesper. Jacobsson. jacobsson.jesper.work@gmail.com

Graphical abstract



Beeswarm plot for over 7000 papers with original perovskite solar cell data. The area of the circles corresponds to the number of citations the papers had obtained by 2022-03-26 and the colour reflects the top device efficiency in each paper. Plot generated with RAWGraphs.

Additional figures

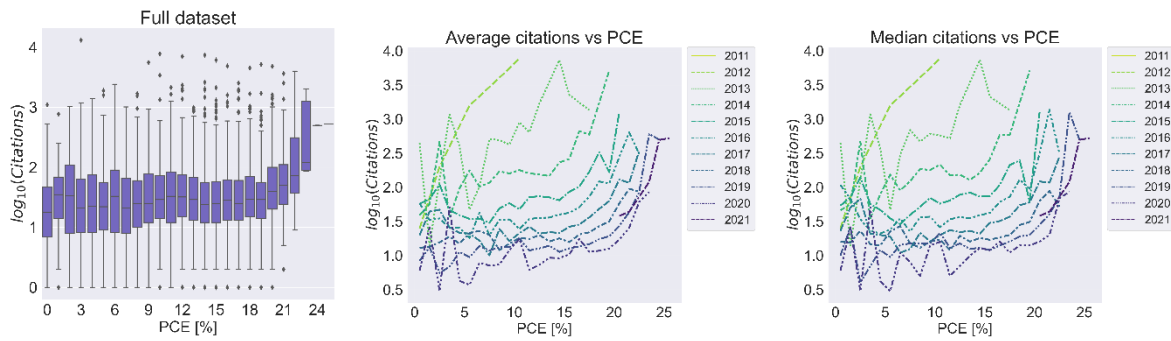


Figure S1. (a) Box plot of the average of the logarithm of the citations for each paper plotted as a function of PCE. The PCE data has been binned in 0.5 % intervals. The data represents the full dataset. Corresponding plots for individual publication years are given in figures below. (b) Average of the log of the number of citations vs. PCE separated by publication year. The PCE data has been binned in 0.5 % intervals and to prevent undefined log-values, 1 has been added to papers with zero citations. (c) Like (b) but median instead of average values.

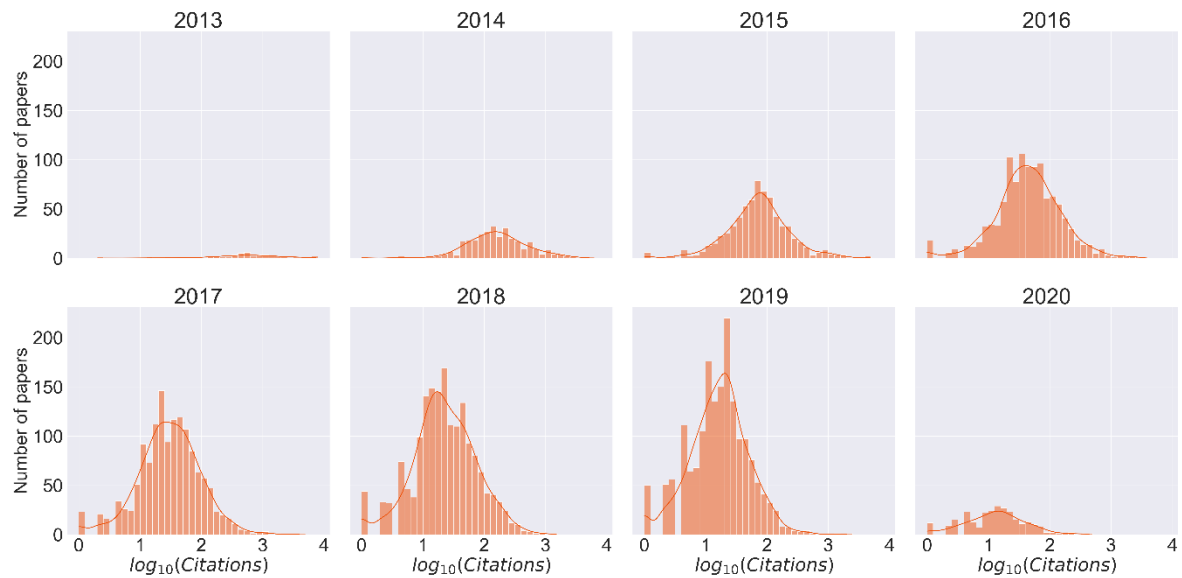


Figure S2. Distribution of citations per paper with one figure for each publication year. The solid line is the associated kernel density estimate.

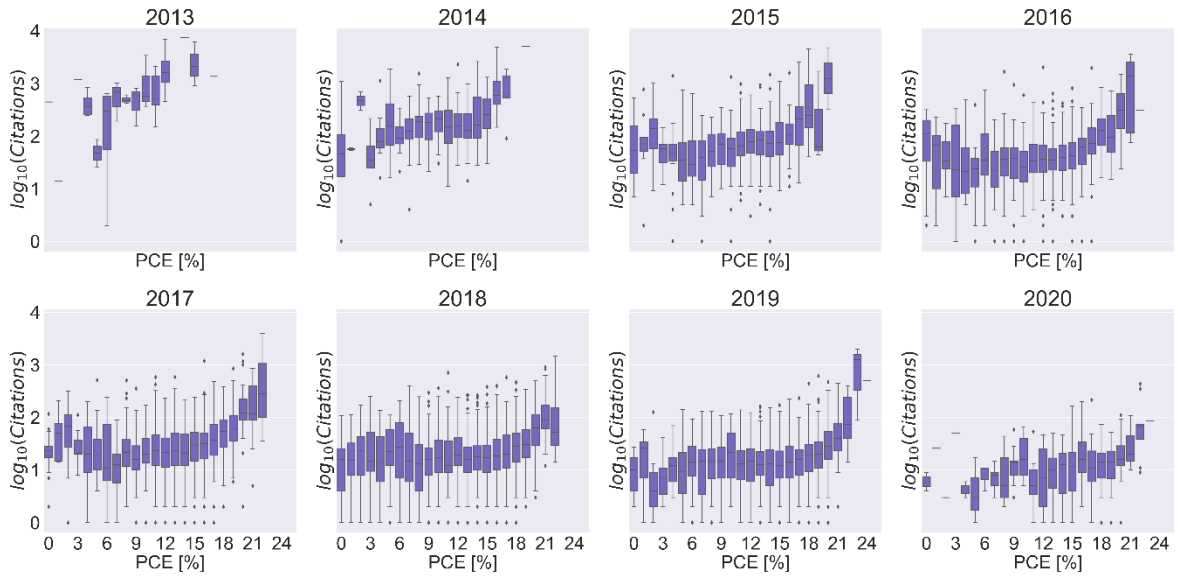


Figure S3. Box plot of the average of the logarithm of the citations for each paper plotted as a function of PCE. The PCE data has been binned in 0.5 % intervals. One figure for each publication year.

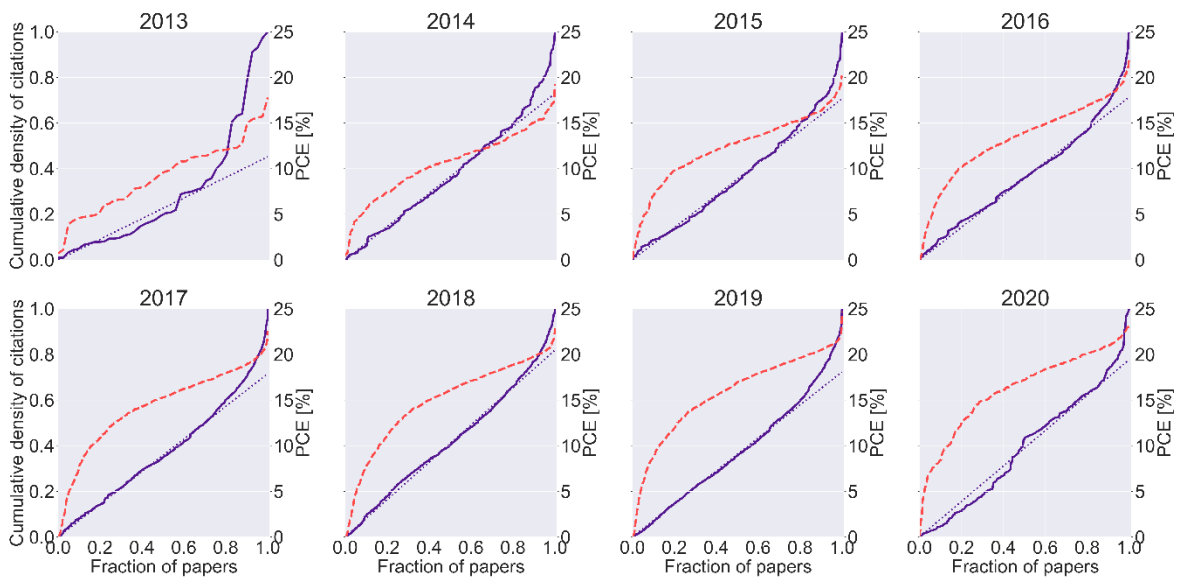


Figure S4. Papers are ordered with respect to the efficiency of the best device reported in them. The cumulative density of citations (i.e., the cumulative sum divided by the total number of citations) is plotted with respect to paper number (divided by the total number of papers to get a scale between 0 and 1). A straight line is fitted to the data between 0 and 0.8. The corresponding device efficiency is given as a red dashed line. One figure for each publication year.



Figure S5. Average of the logarithm of the citations for each paper plotted as a function of PCE. The PCE data has been binned in 0.5 % intervals, and the size of the dots corresponds to the number of papers with that efficiency. Straight lines have been fitted to the data for the top 20 % and bottom 80 % of all papers. One figure for each publication year.

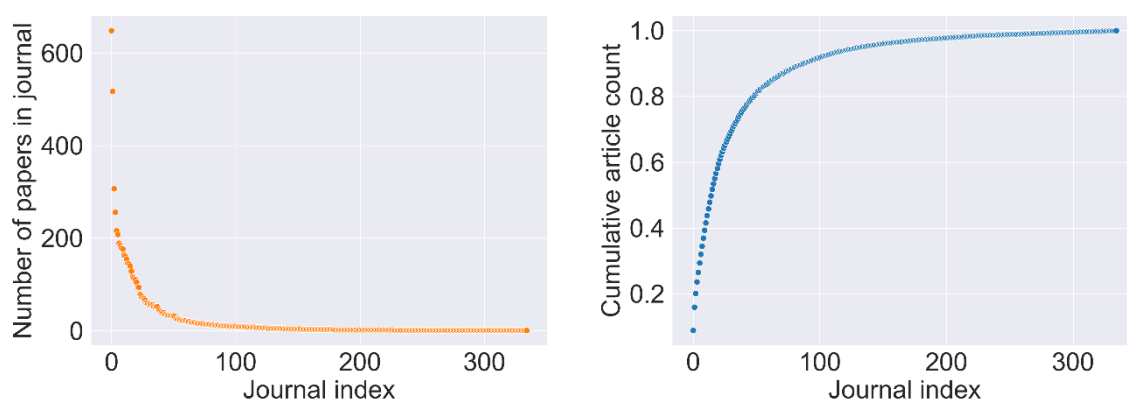


Figure S6. (a) Number of papers per journal. (b) Cumulative count of all perovskite papers with respect to journal. Journals are sorted in order of the number of perovskite papers published in them.

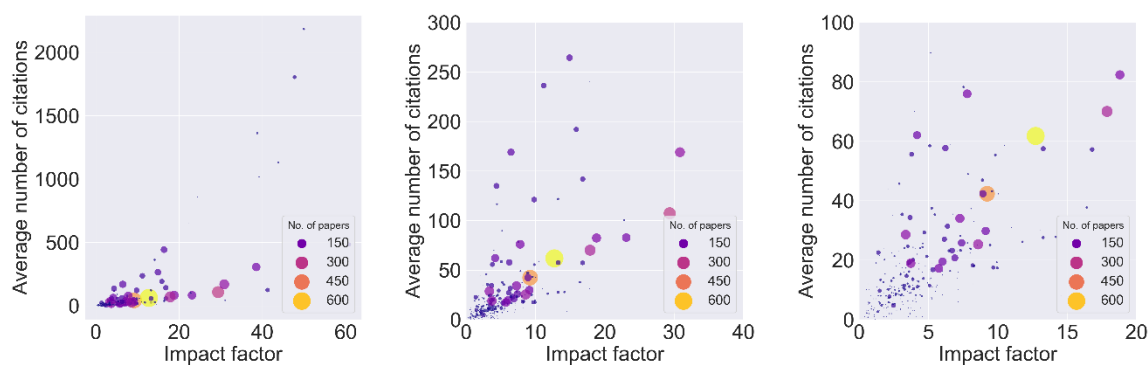


Figure S7. Average number of citations vs. journal impact factor for journals with perovskite papers. Each circle corresponds to one journal, and the size of the circles corresponds to the number of perovskite papers. The figures are the same but with different zoom levels.



Figure S8. Median top-PCE vs. journal impact factor. Each circle corresponds to one journal and the size of the circles corresponds to the median of the citations for all perovskite papers in that journal.

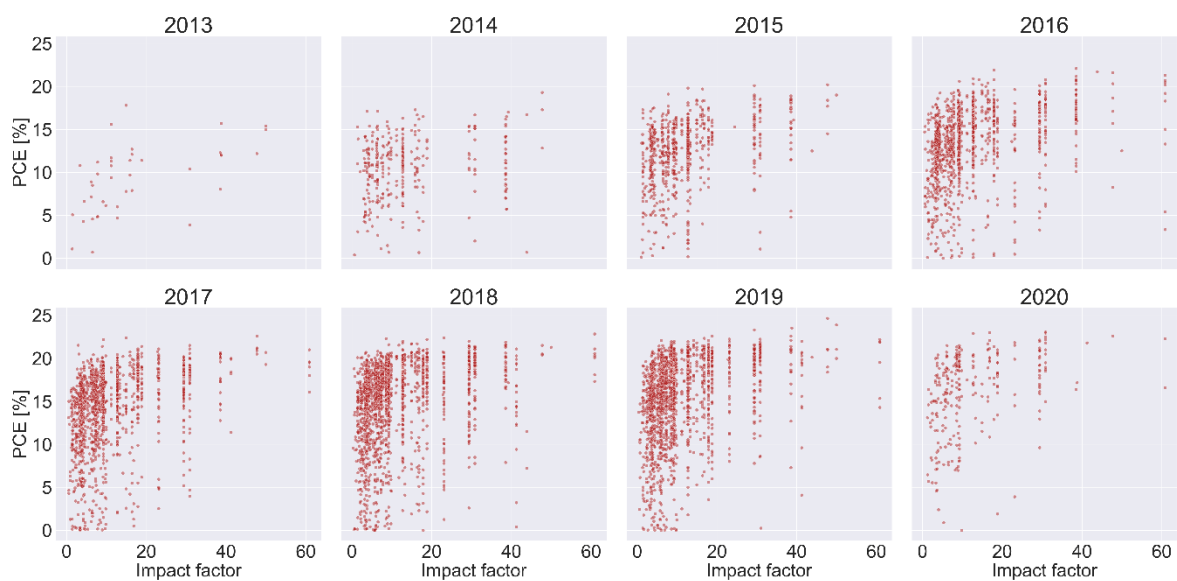


Figure S8. PCE vs. journal impact factor. Each circle corresponds to one paper and the best PCE value reported in that paper.

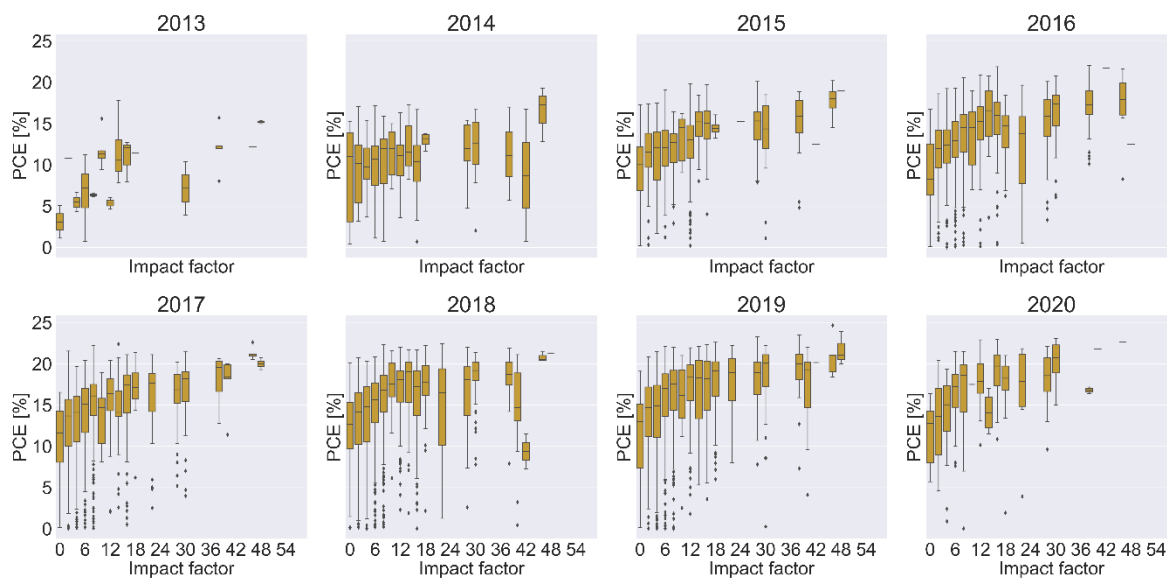


Figure S9. PCE vs. journal impact factor in the form of box plots.

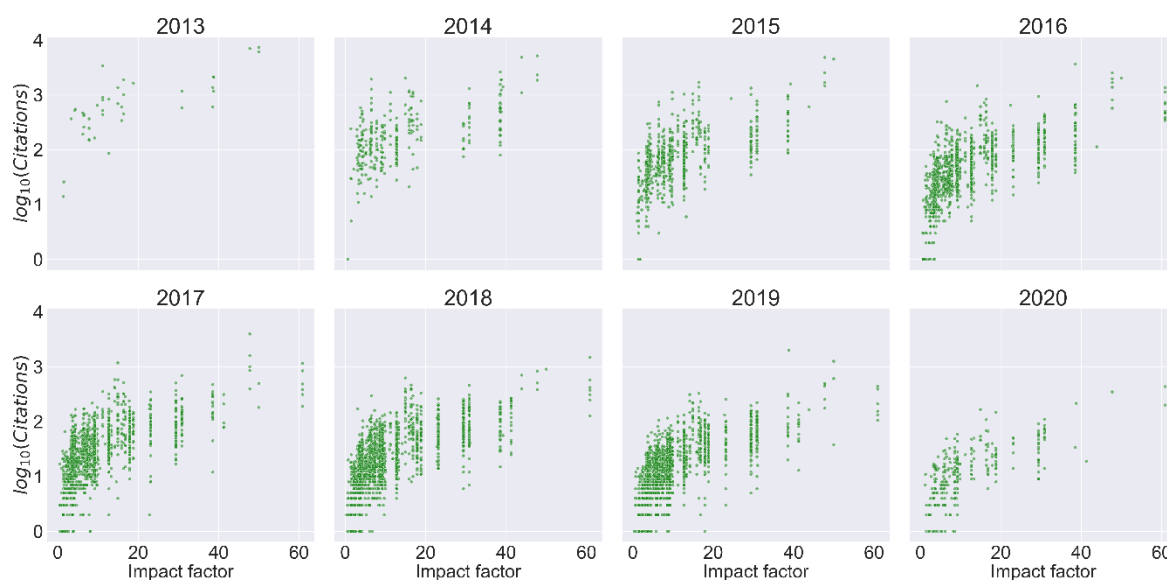


Figure S10. The 10-logarithm of the number of citations vs. journal impact factor. Each circle corresponds to on paper and the best PCE value reported in that paper. To avoid undefined values 1 has been added to each paper with zero citations.

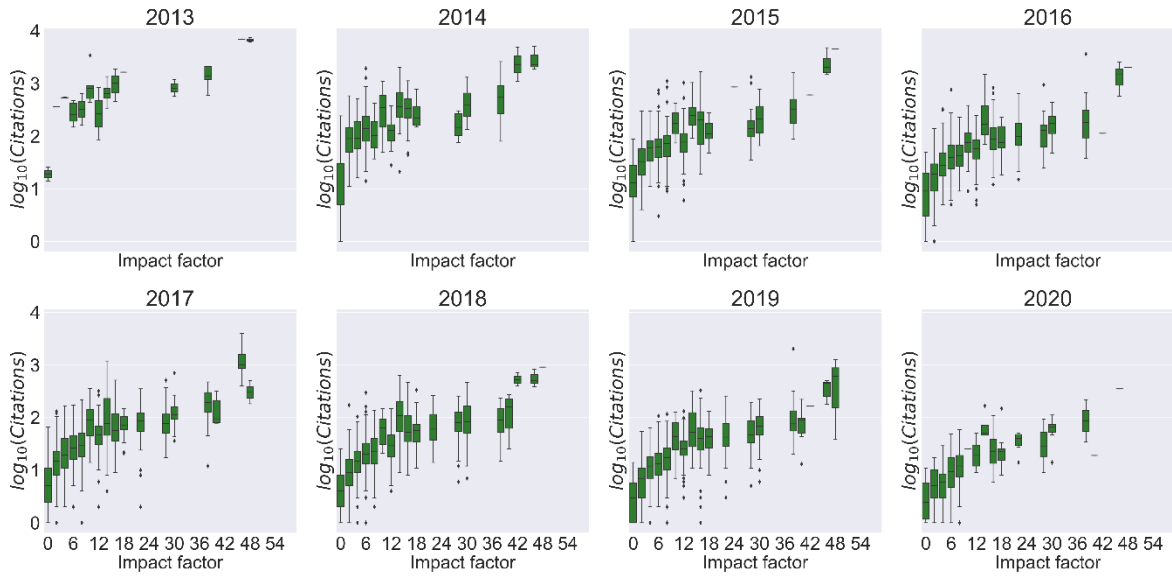


Figure S11. The 10-logarithm of the number of citations vs. journal impact factor in the form of box plots.



Figure S12. Average of the logarithm of the number of citations per journal vs. journal impact factor. Each circle corresponds to one journal.

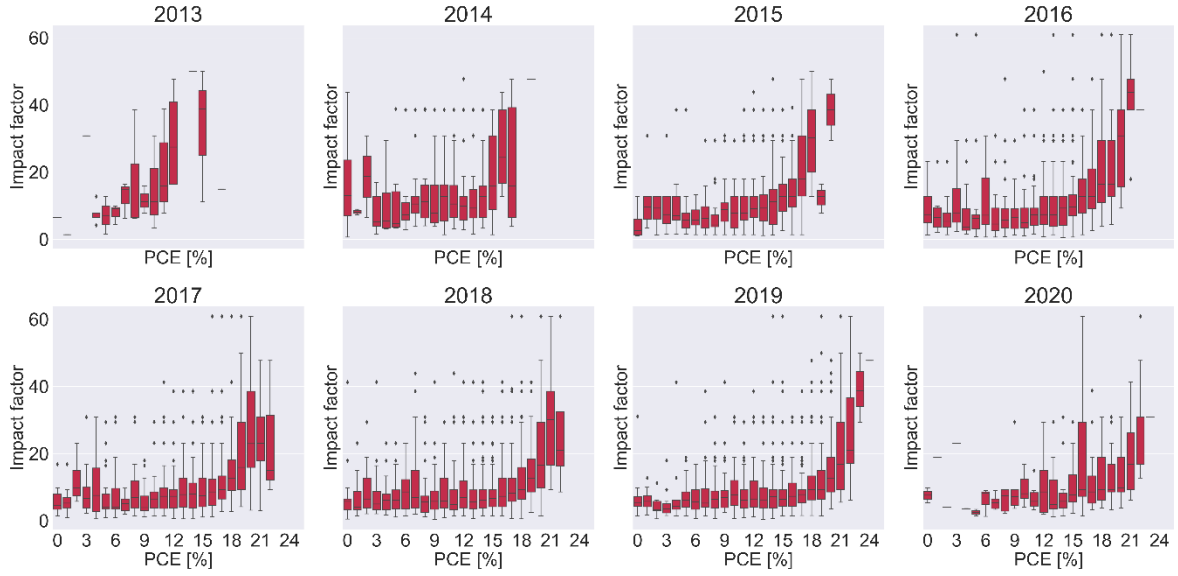


Figure S13. Impact factor vs. PCE-values in the form of box plots. The PCE-values are binned in 0.5 % intervals.

Table S.1 Key metrics from a linear fit of citations with respect to journal impact factor

Year	$d\log(C)/dx$	dC/dx	MEA	R2
2013	0.03	1.07	0.30	0.51
2014	0.02	1.06	0.31	0.30
2015	0.03	1.07	0.30	0.35
2016	0.03	1.07	0.29	0.42
2017	0.03	1.08	0.30	0.38
2018	0.03	1.08	0.30	0.42
2019	0.04	1.08	0.29	0.41
2020	0.03	1.08	0.29	0.45

Table S.2 Key metrics from a linear fit of cell efficiency with respect to journal impact factor

Year	$d\log(C)/dx$	dC/dx	MEA	R2
2013	0.15	1.41	2.66	0.28
2014	0.06	1.14	2.84	0.04
2015	0.13	1.36	2.71	0.11
2016	0.13	1.36	3.09	0.11
2017	0.15	1.42	3.28	0.10
2018	0.13	1.35	3.60	0.07
2019	0.16	1.43	3.53	0.09
2020	0.17	1.47	3.48	0.13