Geoderma 155 (2010) 434–438

Contents lists available at ScienceDirect

Geoderma

journal homepage: www.elsevier.com/locate/geoderma

Individual, country, and journal self-citation in soil science

Budiman Minasny [a], Alfred E. Hartemink [b,*], Alex McBratney [a]

[a] Faculty of Agriculture, Food & Natural Resources, The University of Sydney, McMillan Building A05, NSW 2006, Australia
[b] ISRIC-World Soil Information, P.O. Box 353, 6700 AJ Wageningen, The Netherlands

ARTICLE INFO

Article history:
Received 11 August 2009
Received in revised form 17 November 2009
Accepted 5 December 2009
Available online 1 February 2010

Keywords:
Journal self-citation
Country self-citation
Soil science journals
Pedometrics
Power law
Bibliometrics

ABSTRACT

Self-citation is common practice in most sciences but it differs between disciplines, countries and journals. Here we report on self-citation in soil science. We investigated citations in the major soil science journals and conducted an analysis on a country basis and for the subdiscipline of Pedometrics. It was found that the median rate of individual self-citation was 12%, and ranged from 5 to 60% in 31 soil science journals. A high rate of journal self-citation was accompanied by a high impact factor ranking, but ranking based on the Eigenfactor™ revealed a very different ranking compared to the impact factor score ranking. The distribution of country self-citation rate follows a power law, and a logarithmic function was fitted to the data. Taking into account the logarithmic function, China had high rates of self-citations whereas Egypt, Algeria, Ukraine, and Indonesia have low levels of self-citations. With few exceptions, self-citation rates in soil science are reasonable and comparable to the other biophysical sciences.

1. Introduction

Most authors cite their own work in papers and that seems appropriate as they develop a body of research work and build on previous results and insights. Citing one's own papers or papers from fellow scientists in one's institute or country who work on similar issues is named self-citation. Self-citations account for between 10% and 20% of all references but that differs between scientific disciplines (Hyland, 2003).

Self-citations increase the number of citations and thus the *h* index of an individual scientist (Engqvist and Frommen, 2008) and they may also increase the impact factor of a journal (Hyland, 2003). In a study of citations by Norwegian scientists, Fowler and Asknes (2007) showed that the more one cites oneself, the more one is cited by other scientists. Their analysis suggests that each additional self-citation increases the number of citations from others by about one after 1 year, and by about three after 5 years. In another study, no proof of manipulation of the impact factor through massive use of journal self-citations was found (Andrade et al., 2009).

Little is known about self-citation in soil science. We analysed the trend of self-citations in Pedometrics which is a rapidly growing subdiscipline of soil science. We had the impression that the self-citation rates differed between countries and then investigated self-citations by countries and for different soil science journals; self-citation here means a paper citing other papers from the same country or the same journal.

2. Self-citation in Pedometrics

It is difficult (practically impossible) to obtain self-citation rates for all soil science papers. As a representation, we investigated the self-citation rate in a subdiscipline of soil science: Pedometrics. We analysed and manually counted papers from Pedometrics Special Issues that have been published in *Geoderma* between 1994 to 2007. The Pedometric special issues are published by *Geoderma* every 2 years. There have been nine scientific Pedometrics symposia starting in 1992 (Table 1) and there were a total of 105 papers. We manually counted the number of references and self-citations for each paper. Self-citation means the number of references cited in the paper that are written by any of the paper's authors.

The number of references in research papers (review papers excluded) is slightly skewed to the left, with a minimum of 7, a maximum of 91, a median of 28, and an interquartile range of 19 to 38 references. Pedometrics papers have on average 30 references. There is a slight increase in the number references over time (Fig. 1), possibly due to the advances in electronic and online journals, making it easier to find more relevant references. Also the field of Pedometrics is expanding and it becomes necessary to refer to more work. In the beginning, there were simply less papers that could be cited.

Fig. 2 shows the histogram of the percentage of self-citations for Pedometrics papers. Papers in Pedometrics special issues of *Geoderma* have a self-citation rate of around 14%. Review papers have larger self-

* Corresponding author. Tel.: +31 317 471713.
E-mail address: alfred.hartemink@wur.nl (A.E. Hartemink).

0016-7061/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.geoderma.2009.12.003

Table 1
Special issues of *Geoderma* and the year published and number of papers.

Symposium	Location	Year published	Number of papers
PM1992	Wageningen, The Netherlands	March 1994	20
Fuzzy Sets	St. Louis, USA	June 1997	12
PM1997	Madison, USA	April 1999	7
PM1998	Montpellier, France	Sept. 2000	14
PM1999	Sydney, Australia	Sept. 2001	11
PM2001	Gent, Belgium	March 2003	8
PM2003	Reading, UK	Oct. 2005	14
PM2005	Florida, USA	Aug. 2007	10
PM2007	Tübingen, Germany	2010	9

citation rates (interquartile range of 20 to 30%) compared to research papers (interquartile range of 10 to 18%), based on the analysis of seven review papers. Although there appears a slight increase in self-citation with the number of authors over time, the trend is not statistically significant. There is also no observable trend of self-citation rate with time.

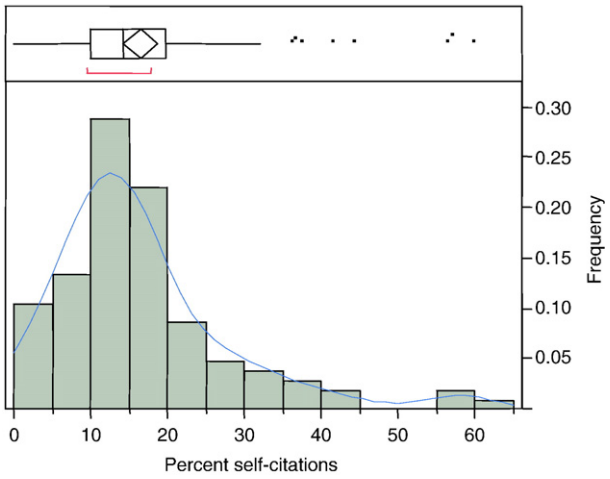

Fig. 2. Histogram of percentage of self-citations in Pedometrics papers (1992–2007).

3. Country self-citations

We used the soil science publication data from SCImago for the period 1996–2007. The SCImago Journal & Country Rank (www.scimagojr.com) is a portal that includes journals and country scientific indicators developed from information in the Scopus database (Elsevier Science). These indicators can be used to assess and analyse scientific domains. Country self-citation means the percentages of the citations received by the papers which come from the same country as from which the papers were published. It includes authors citing papers from fellow scientists of their own country. Fig. 3 shows the number of papers produced between 1996 and 2007 and the percentage of country self-citations from 170 countries. We fit a log model to the data:

$$\text{Percent self-citation} = 9.8 * \text{Log}_{10}(\text{number of papers}).$$

The *x*-axis in Fig. 3 is on a logarithmic scale as the distribution of citations usually follows a power law (Redner, 1998). Initially with a small amount of papers, the citation increases slowly as only people in the same field (or country) cite the papers. As the number of paper increases, more people will notice the papers and the citation rates increases.

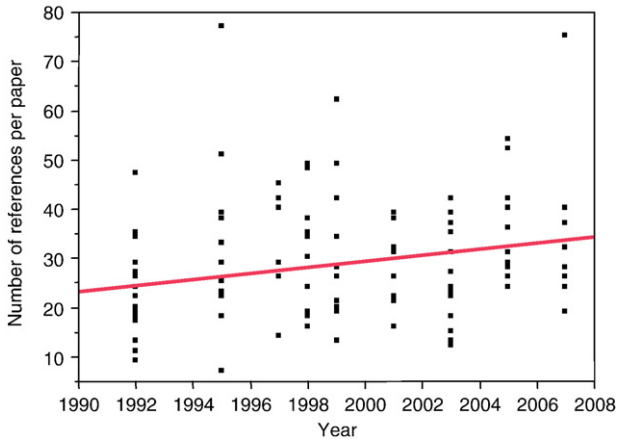

Fig. 1. Number of references in Pedometrics special issues of *Geoderma* from 1992 to 2007. The line represents a linear regression.

Countries with the highest number of self-citations are China (63%) and the USA (48%). The trend seems to be that with every tenfold increase in the number of papers, there is a 10% increase in the number of self-citations. So the more papers a country produce, the more likely it will refer to papers of its own nation. This is possibly because the more papers a country produced, there is a higher chance that a person from that country will cite more work from its own country — soil scientists in countries with a large body of work are more inclined to cite papers from their own fellow countrymen and women. Smaller countries have fewer scientists and papers and hence cite more papers from other countries.

The residuals of the log model can be plotted and the countries that depart most from the trend or zero residuals can be viewed (Fig. 4). Some countries tend to over-cite themselves, but there are also countries with low self-citations. The USA seems to have a high country self-citation rate. According to the trend line it should have a self-citation rate of 41% and the residual does not depart far from zero. This is because the USA produces a large body scientific work (23% of the soil science papers). Meanwhile China, Serbia, Libya have high

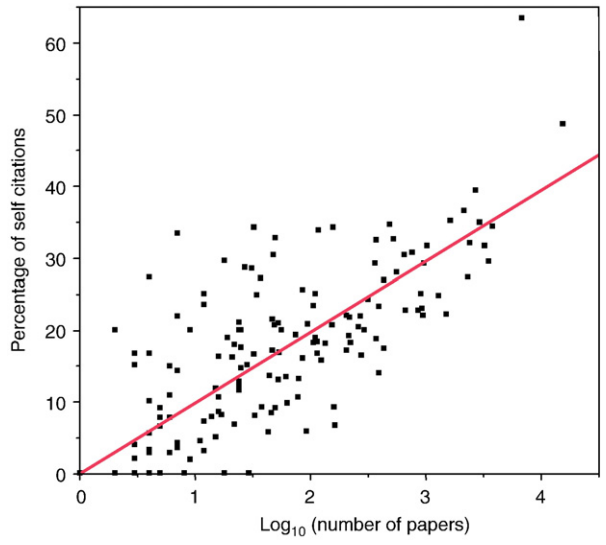

Fig. 3. Log (number of papers) produced by 170 countries in the area of soil science 1996–2007 and its relationship with percent of self-citations. The line is a log–linear model: Percentage self-citation = 9.8 Log_{10}(number of papers). Data from SCImago.

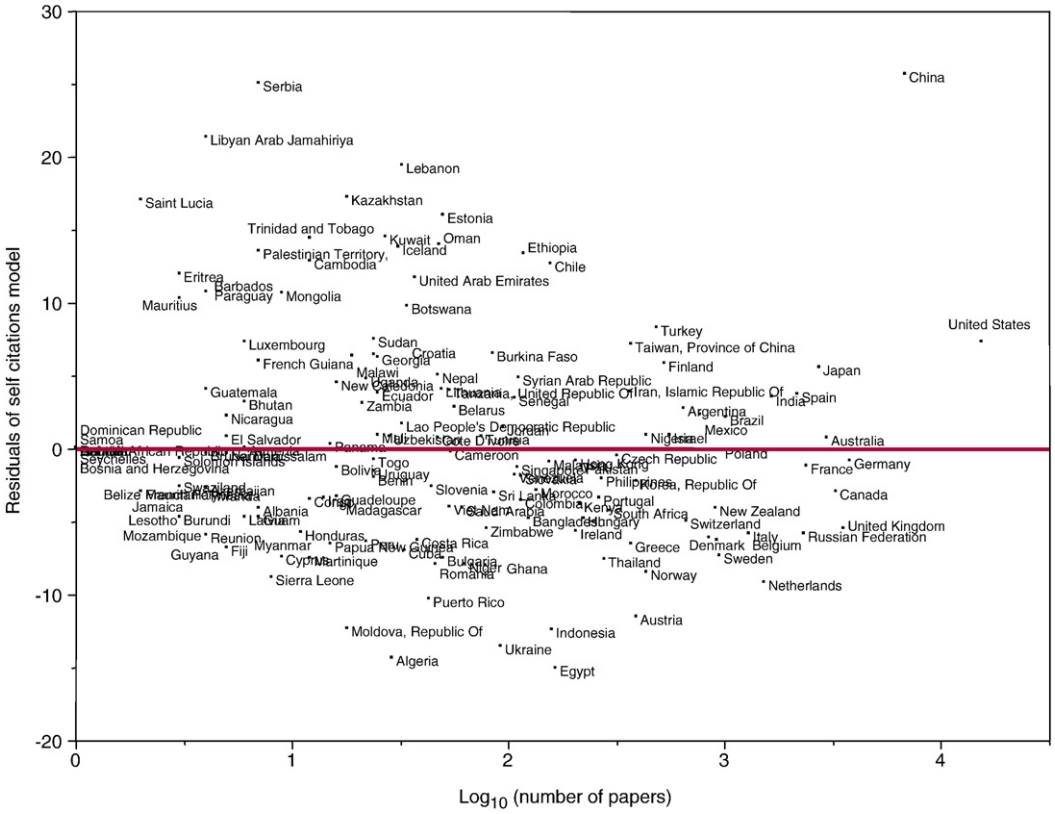

Fig. 4. Residuals of the regression line from Fig. 3.

residuals, meaning that they tend to over-cite themselves. Egypt, Algeria, Ukraine, and Indonesia have low levels of self-citations.

The trend line (Fig. 3) gives the likely country self-citation rate for a paper. For example, a paper from Australia would have on average 34% of self-citations, the Netherlands 31% whereas self-citation in soil science papers from the UK and Germany is about 35%. A list of the top 20 countries of soil science paper producers is given in Table 2, with the likely self-citation rate.

Table 2
The top 20 countries in soil science publications during the period of 1996–2007.

Country	No. citable publications	No. citations	Actual self-citations (%)	Regression model of self-citations (%)
United States	15,452	138,272	48.5	41.3
China	6773	17,697	63.4	37.7
Germany	3762	41,059	34.3	35.2
United Kingdom	3519	47,431	29.4	34.9
Canada	3246	30,659	31.6	34.6
Australia	2953	28,965	34.9	34.2
Japan	2727	14,414	39.3	33.8
France	2391	27,150	32.0	33.3
Russian Federation	2318	3848	27.3	33.1
Spain	2153	16,659	36.5	32.8
India	1643	7431	35.2	31.7
The Netherlands	1519	20,589	22.2	31.3
Italy	1298	11,088	24.7	30.7
Brazil	1013	7886	31.7	29.6
Poland	958	3249	29.2	29.4
Sweden	945	12,488	22.0	29.3
Belgium	922	9298	22.9	29.2
New Zealand	902	10,861	25.0	29.1
Denmark	849	11,384	22.7	28.8
Mexico	765	2514	30.7	28.4

4. Journal self-citations

Journal self-citation here means you cite papers that are from the same journal. Table 3 shows the 2008 Citation Reports from Thomson–Reuters (previously ISI) for major soil science journals. The impact factor in 2008 is calculated as the number of citations in 2008 to papers published in 2006 and 2007 divided by the number of papers published in 2006 and 2007 (Garfield, 2006). The percent of self-citation in Table 3 refers to the amount of self-citation that is used in the impact factor calculation. There is a great difference between the soil science journals — the percentage of self-citations ranges from 5 to 60%.

The distribution of self-citations (Fig. 5) is skewed by three outliers (*Journal of Soils and Sediments, Agrochimica*, and *Revista Brasileira de Ciencia do Solo*). The median of self-citations is 12%, which seems to be about a normal self-citation rate for a soil science journal. By comparison, the self-citation rate in *Nature* and *Science* is 1%.

There is another metric called the Eigenfactor™ score that counters this problem. The Eigenfactor™ score ranks the influence of journals in the same way as Google's PageRank algorithm ranks the influence of web pages. Journals are considered to be influential if they are cited often by other influential journals (see www.eigenfactor.org for more details).

The rank according to impact factor and Eigenfactor score is plotted in Fig. 6. Most journals are close to the 1:1 line except for the *Journal of Soils and Sediments* which indicates that its self-citation favours its high impact factor. *Soil Biology and Biochemistry* ranks first for both scoring methods. There are journals which have a lower impact factor rank, but higher Eigenfactor rank. For example, *Soil Science* and the *Australian Journal of Soil Research* have low impact factor rank, but these journals are being cited by more influential journals. Meanwhile, *Soil Use and Management* has a higher impact factor, but may not be cited by more influential journals.

Table 3
Impact factor, Eigenfactor score and percentage self-citations of major soil science journals (2008 data).

Journal	Rank by Impact factor	Impact factor	Rank by Eigenfactor score	Eigenfactor score	% self-citations (used in Impact factor calculation)
Soil Biology and Biochemistry	1	2.926	1	0.03265	18
Journal of Soil and Sediment	2	2.797	25	0.00164	42
Applied Soil Ecology	3	2.247	8	0.00838	11
European Journal of Soil Science	4	2.24	6	0.01024	8
Soil Science Society of America Journal	5	2.207	3	0.02381	12
Geoderma	6	2.068	4	0.01978	15
Plant and Soil	7	1.998	2	0.02721	12
Soil Use and Management	8	1.895	17	0.00409	9
Catena	9	1.874	9	0.00773	10
Soil and Tillage Research	10	1.695	5	0.01136	10
Pedobiologia	11	1.451	16	0.0041	5
Biology and Fertility of Soils	12	1.446	10	0.00707	10
Vadose Zone Journal	13	1.441	7	0.0084	28
Journal of Plant Nutrition and Soil Science	14	1.284	12	0.00504	11
Nutrient Cycling in the Agroecosystems	15	1.282	11	0.00509	8
Land Degradation and Development	16	1.245	22	0.00239	16
Clays and Clay Minerals	17	1.171	15	0.00411	12
Soil Science and Plant Nutrition	18	1.152	20	0.00294	29
Journal of Soil and Water Conservation	19	1.121	19	0.0032	22
Soil Science	20	1.037	13	0.00456	7
Canadian Journal of Soil Science	21	1.023	21	0.00294	14
European Journal of Soil Biology	22	0.888	24	0.00189	12
Pedosphere	23	0.865	23	0.00203	8
Australian Journal of Soil Research	24	0.856	14	0.00438	20
Revista Brasileira de Ciencia do Solo	25	0.66	26	0.00119	60
Compost Science and Utilization	26	0.638	27	0.00115	25
Acta Agriculturae Scandinavica — Section B Soil and Plant Science	27	0.407	28	0.00067	8
Communications in Soil Science and Plant Analysis	28	0.357	18	0.00401	7
Arid Land Research and Management	29	0.348	29	0.00064	12
Agrochimica	30	0.179	31	0.00024	40
Eurasian Soil Science	31	0.149	30	0.00048	53

5. Discussion and conclusions

Hyland (2003) found that self-citation is higher in the "hard" sciences (biology, engineering and physics), where it is over 12% of all references, compared to only 4% in sociology, philosophy, linguistics, or marketing. For all soil science journals we found a 12% self-citation rate but presumably it will differ between the subdisciplines. The soil science subdiscipline Pedometrics is 14% — it is a young subdiscipline and initially dominated by a few people and a few seminal papers that are often cited. We think that the number of self-citations will decrease over the years when Pedometrics further matures and the number of Pedometricians and papers increase.

Although we have not looked at individual soil scientists, we found large differences between self-citation rates for the major soil science journals. High rates of self-citations can influence the impact factor of a journal. Andrade et al. (2009) found no proof for manipulation of the impact factor through self-citation. It seems that the high impact factor of the *Journal of Soils and Sediments* (ranked second in 2008) is an exception — 42% of the 207 citations used in the 2008 impact factor calculation are from the journal itself. *Agrochimica* and *Revista Brasileira de Ciencia do Solo* have high self-citation rate (40% and 60%, respectively) but do not have a high impact factor. In *Revista Brasileira de Ciencia do Solo*, the journal has produced a larger number of papers (265 papers) and has a relatively low number of citations (175). This is probably because it is one of the few soil science journals in Portuguese, it has a high self-citation rate but that is not directed to influence the impact factor. In order to boost the impact factor with self-citations, the number of papers should be low (the denominator in the impact factor formula) and the number of self-citations (the numerator) should be high.

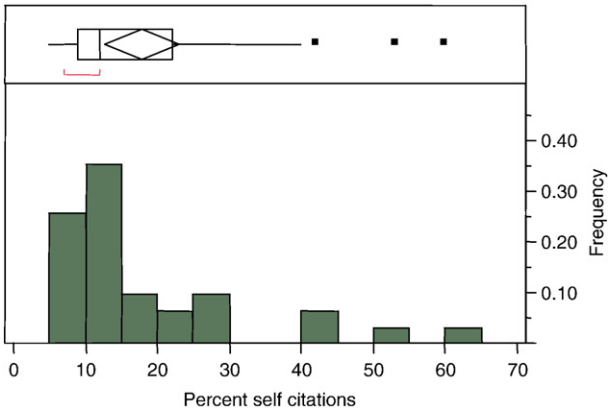

Fig. 5. Distribution of percentage of self-citations from soil science journals used in the impact factor calculation of 2008.

Another trick to increase citations and impact factor is self-citation through "editorial material" (González and Campanario, 2007; Epstein, 2007). That means the editor of the journal cites papers in own journal in the editorial material, and this is counted in the impact factor calculation. In the impact factor calculation, editorial materials are not taken into account in the denominator (the number of papers), however self-citations contribute positively to the numerator (number of citations).

Are large self-citation rates narcissistic or manifestations of laziness or extreme conviction? As summarized by Hyland (2003) "the factors which motivate writers to cite their own work are doubtless varied and complex, involving psychological factors influenced by the individual writer's confidence, experience and self-esteem." Here some data are presented that underpin the impression that high rates of self-citation are quite accepted in some countries and in some journals. Self-citation is a contagious issue. It can increase your *h* index, and also your country's pride and

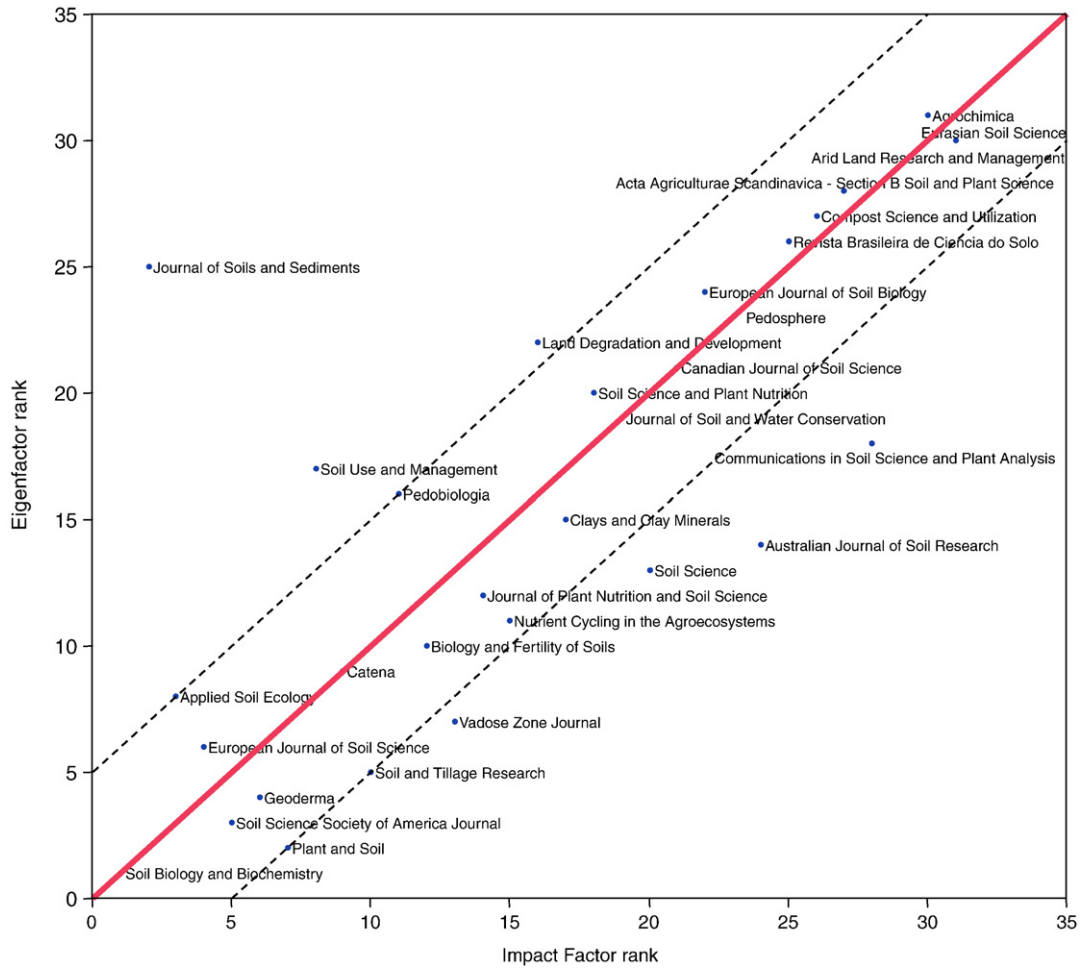

Fig. 6. Soil Science journals' rank according to impact factor and Eigenfactor score (2008 data). The thick line represents a 1:1 line, the outer lines represent a score different of 5 between the two rankings.

your preferred journal's impact factor. There is no guideline and certainly not a penalty for high rates of self-citations. However, we are convinced that reality–and certainly time–will catch up with those that fiddle with good scientific practice and reasonable degrees of self-citation.

References

Andrade, A., González-Jonte, R., Campanario, J.M., 2009. Journals that increase their impact factor at least fourfold in a few years: the role of journal self-citations. Scientometrics 80, 515–528.

González, L., Campanario, J.M., 2007. Structure of the impact factor of journals included in the social sciences citation index: citations from documents labeled "Editorial Material". Journal of the American Society for Information Science and Technology 58, 252–262.

Engqvist, L., Frommen, J.G., 2008. The h-index and self-citations. Trends in Ecology & Evolution 23, 250–252.

Epstein, D., 2007. Impact factor manipulation. The Write Stuff 16, 133–134.

Fowler, J.H., Asknes, D.W., 2007. Does self-citation pay? Scientometrics 72, 427–437.

Garfield, E., 2006. The history and meaning of the journal impact factor. Journal of the American Medical Association 295, 90–93.

Hyland, K., 2003. Self-citation and self-reference: credibility and promotion in academic publication. Journal of the American Society for Information Science and Technology 54, 251–259.

Redner, S., 1998. How popular is your paper? An empirical study of the citation distribution. European Physics Journal B 4, 131–134.