Journal of Informetrics 7 (2013) 357–370

Contents lists available at SciVerse ScienceDirect

Journal of Informetrics

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

Correlation between Journal Impact Factor and Citation Performance: An experimental study

Ugo Finardi[a,b,*]

[a] CNR-CERIS, National Research Council of Italy, Institute for Economic Research on Firms and Growth, via Real Collegio, 30, I-10024 Moncalieri (TO), Italy
[b] University of Torino, Department of Chemistry and NIS-Centre of Excellence, via P. Giuria, 7, I-10125 Torino, Italy

ARTICLE INFO

Article history:
Received 24 September 2012
Received in revised form
21 December 2012
Accepted 21 December 2012
Available online 9 February 2013

Keywords:
Garfield
Journal Impact Factor
Article citations
Correlation coefficients
Time evolution
Impact factor misuses

ABSTRACT

Since its introduction, the Journal Impact Factor has probably been the most extensively adopted bibliometric indicator. Notwithstanding its well-known strengths and limits, it is still widely misused as a tool for evaluation, well beyond the purposes it was intended for. In order to shed further light on its nature, the present work studies how the correlation between the Journal Impact Factor and the (time-weighed) article Mean Received Citations (intended as a measure of journal performance) has evolved through time. It focuses on a sample of hard sciences and social sciences journals from the 1999 to 2010 time period. Correlation coefficients (Pearson's Coefficients as well as Spearman's Coefficients and Kendall's τ_α) are calculated and then tested against several null hypotheses. The results show that in most cases Journal Impact Factors and their yearly variations do not display a strong correlation with citedness. Differences also exist among scientific areas.

1. Introduction

Since it was first mentioned by Garfield (1955), the Journal Impact Factor (JIF from now on) has probably been the most widespread and relevant metric used to evaluate scientific literature.

It is well known that the JIF for year *i* is defined as:

$$JIF_i = \frac{CIT_i(i-1) + CIT_i(i-2)}{JART_{i-1} + JART_{i-2}}$$

where $CIT_i(i-1)$ and $CIT_i(i-2)$ are the citations received in year *i* by *citable items* (of the specific Journal) published during the two previous years, while $JART_{i-1}$ and $JART_{i-2}$ are the number of *citable items* published in those same two years (Garfield & Sher, 1963, 1966; Garfield, 1972; see also Bensman, 2007a; Garfield, 1994).

The first Science Citation Index appeared in 1963 (Pendlebury, 2009), and the JIF has been published by the Institute for Scientific Information® (ISI®), now Thomson Reuters®, since 1975 (Garfield, 1994). JIFs and numbers of received citations can be retrieved via the Web of Knowledge website.

Since its appearance, and particularly in more recent years, a growing stream of scientific literature has discussed the features and applications, as well as limits and misuses, of the JIF, witnessing its great diffusion in most scientific fields. Garfield himself has discussed the misuses of the JIF, stating for instance: "I expected it to be used constructively while

* Corresponding author at: CNR-CERIS, National Research Council of Italy, Institute for Economic Research on Firms and Growth, via Real Collegio, 30, I-10024 Moncalieri (TO), Italy. Tel.: +39 011 6824923.
E-mail addresses: u.finardi@ceris.cnr.it, ugo.finardi@unito.it, ugo.finardi@cnr.it

1751-1577/$ – see front matter
http://dx.doi.org/10.1016/j.joi.2012.12.004

recognizing that in the wrong hands it might be abused" (Garfield, 2005, p. 1) and "it is one thing to use impact factors to compare journals and quite another to use them to compare authors" (Garfield, 2005, p. 4); "The source of much anxiety about Journal Impact Factors comes from their misuse in evaluating individuals. [. . .] I have always warned against this use" (Garfield, 1997, p. 1). In spite of this, "The journal impact factor [. . .] has nowadays become the bibliometric construct most widely used for evaluation in the scholarly and publishing community" (Moed et al., 2012, p. 368).

The present work aims at contributing to the debate on the features of the JIF, and on its uses and misuses. In particular, it tackles a specific research question: is there a correlation (and, if so, how strong is it?) between Impact Factors and citations received by a journal in a given year and their time evolution? JIFs are very often used as a measure of the quality of journal articles published in the year they refer to, in order to evaluate scientific literature (the production of Research Institutions or individuals, or even single works). Thus, it may prove important to check whether this use (which should be considered a misuse) has some meaning or, instead, whether more evidence suggesting its avoidance can be gathered.

If it were meaningful to use the JIF of the year of publication to evaluate that year's articles, one would expect a very strong (linear or almost linear) correlation between how the JIF and Mean Received Citations (MRC from now on) evolve over time, thus justifying this use. Otherwise, it would be difficult to infer that the articles of a specific journal in a specific year are more/less valuable than those of the same journal in another year, or those of another journal in the same year. The MRC, which will be analytically defined in Section 3, is the mean number of citations received by articles published in the year, normalised by their age.

In particular, this work explores the correlation existing between the evolution of the JIF and of the corresponding MRC (for year and journal) focusing on a sample of journals from the 1999 to 2010 time period. Correlation is thus considered dynamically, taking into account the evolution of the two indicators over time. Moreover, the MRC values are normalised for the age of the articles and thus made time-independent, just like the JIF values are. As the Theoretical Framework section will explain, measures of citations are generally normalised across scientific sectors, in order to be as free from field effects as possible.

The MRC and JIF have been obtained (from the in-house version of the Web of Knowledge at the University of Turin, Italy) for a sample of journals belonging to two ISI® Subject Categories ("Chemistry, Multidisciplinary" and "Management"), chosen for being rather general in their scope. The sample includes a total of 87,766 articles. Three types of correlation coefficients have been calculated from the obtained data: Pearson's Correlation Coefficients, Spearman's Rank-Order Correlation Coefficients, and Kendall's $\boldsymbol{\tau_\alpha}$. The use of the last two (non parametric) coefficients also aims at overcoming the problems that may arise using parametric statistics in skewed distributions, such as those of citations.[1] In order to perform a further analysis, the obtained coefficients have then been tested against several null hypotheses, covering the range from linear (1) correlation to 0 correlation coefficient.

The paper is organised as follows: Section 2 illustrates the theoretical framework, with specific regard to the most recent literature on the features of the JIF; Section 3 describes the methodology of the work; Section 4 reports the obtained results, while Section 5 presents their discussion and some conclusions.

2. Theoretical framework

In order to better describe the experimental activity, this section presents the results of some previous works on the subject, also showing the lack of literature on the specific topic discussed in this article. Reference is also made to the past literature on the use of means of citations in order to better illustrate how they are used in the present work.

Garfield (2005, 2006) offers a historical outline of the JIF and of why and how it was introduced, also discussing several of its features. Archambault and Larivière (2009) examine how measures of journal impact evolved, leading to the JIF; moreover, they mention some of its misuses (such as comparing the JIF across different scientific fields) and criticise it, affirming that "For all of the reasons presented in this paper, the indicators presented in the Journal Citation Report (JCR) cannot be considered robust objective measures of the worth of all journals published internationally" (p. 644).

Besides Garfield's comments mentioned above, there have been several criticisms and warnings against misusing the JIF, also in non-bibliometry-specific journals, thus confirming the importance of this topic.

Seglen (1997) discusses whether or not the JIF is representative of individual journal articles, also listing various problems associated with its use. His findings lead him to conclude that "the journal cannot in any way be taken as representative of the article. Even if it could, the Journal Impact Factor would still be far from being a quality indicator" (p. 502). The experimental activities presented in this work can provide further evidence to support this opinion.

In a previous paper, Seglen (1994) studies the connection between article Citedness and journal impact in a database of articles published by a set of 16 scientists (for a total of about 1000 articles). Citations present a highly skewed distribution: "Thus [. . .] journal impact is unsuitable as a proxy measure of article Citedness" (p. 9).

Distribution of citations in journal articles has also been studied by Vieira and Gomes (2010) and by Mutz and Daniel (2012). The former article looks at the distribution of citations in a specific classification in the Web of Science for a specific year (taking four Subject categories into account). Its findings show that "the correlations found for the means of the

[1] See, for instance, Albarrán and Ruiz-Castillo (2011) for a discussion on skewed distribution of citations.

citations of large sets of documents do not allow direct conclusions for individual documents" (p. 11). Moreover, there is direct dependence between citation counts (mean cites per article) and other properties: the most meaningful in the present context is the JIF. In our paper, however, citations are studied in their evolution over time, rather than in a specific year.

Mutz and Daniel (2012) propose solutions to two core problems existing in JIF computation: effects of skewed citation distribution and biasing factors (such as citable document types). They conclude their analysis by suggesting a new methodological approach to the comparison of sets of scientific literature.

Warnings about the problems that may arise from using citation indicators in the evaluation of research and journals are also voiced by Leydesdorff (2006, 2008). According to his experimental work, the classification of journals by subject categories cannot be performed unambiguously taking interrelation of citations into account; also non-ISI journals have a relevant role in citation patterns. On the basis of his findings, he states that: "impact factors provide a summary statistics that conveniently allows for the ranking of journals. However, these statistics are based on the means of a highly skewed distribution" (p. 284, 2008).

Pendlebury (2009) describes the history, strengths and limitations of – and alternatives to – the JIF, and lists the "Ten Commandments" of citation analysis; he then concludes that "the consequences of such misuse can be profound" and that "numbers can be dangerous because they have the appearance of being authoritative" (p. 9). The misbehaviour of editors is studied instead by Wilhite and Fong (2012), who use the results of a survey to suggest that authors are often coerced to cite articles published in the journal to which they are submitting their work in order to increase its citation rate. On the other hand, an Editorial by the PLoS Medicine Editors (2006) criticises "Thomson Scientific, the sole arbiter of the impact factor game" (p. 0707), concluding that "science is currently rated by a process that is itself unscientific, subjective, and secretive".

Bensman (2007b) discusses the features of the JIF and calculates correlation coefficients for JIF and SCI Total citations for Chemistry journals in 1993, as well as for other measures of journal importance. His findings show that "total citations are a better holistic measure of journal importance than the impact factor" (p. 62).

In a recent paper, Vanclay (2011)[2] criticises the specific way in which the JIF is applied and calculated by ISI-Thomson Reuters. He states that "a major overhaul is warranted, and [. . .] users should cease using the Thomson-Reuters Impact Factor until Thomson-Reuters has addressed these weaknesses" (p. 430). His article has stirred a wide debate among scholars. For instance, Campanario (2012) comments on some of its points, also affirming that "To shed light on this topic, an initial task could be to study the correlations between JIF and other performance measures" (p. 294). Indeed, the present article attempts to fill this knowledge gap. Another widely debated feature of the JIF is its 2-year citations window. This matter is addressed more extensively by van Leeuwen (2012), who – analysing aggregated data concerning five subject categories – shows that impact in the short term is representative of citation impact in the years to come.

Ingwersen (2012) suggests the use of a *diachronic* JIF, obtained by calculating the number of citations received by journal articles of the same year for which the JIF is calculated *after* they are published. The present work adds further evidence to support this approach. Also Balaban (2012) suggests the use of the JIF coupled with the number of received citations in order to build a citation index that can be applied individually.

Other, more direct, responses to Vanclay's paper (Bensman, 2012; Moed et al., 2012; Pendlebury & Adams, 2012; Rousseau, 2012; Vinkler, 2012a) either partly support or criticise his statements. Most authors support the use of indicators – and particularly of the JIF – *in an informed way* in the assessment of journals, and also tend to distinguish between "Garfield's JIF" and its modern applications and possible misuses.

A topic which is particularly relevant to the experimental activity carried out in the present work is that of means of citations, which have been intensively studied in recent years. For instance, Albarrán et al. (2010) evaluate the U.S. vs. E.U. scientific production by using the "normalised mean citation rate per article" (p. 336 and following), MCR, calculated as a ratio of fractions of received citations and number of articles. In their analysis "scientific performance is identified with the impact that journal articles have through the citations they receive" (p. 329).

A vast literature deals with the production of field-independent, rather than time-independent, indicators. For instance, while performing a normalising exercise, Wu (2013) cites a large amount of past works on the topic. Furthermore, Herranz and Ruiz-Castillo (2012) investigate the citation impact of world geographical areas and use normalised indicators (such as the crown indicator and the Mean Normalised Citation Score) with sub-field normalisation procedures.

More generally, several "relative scientometric indicators" (Vinkler, 2003) or "relative impact indicators" (Vinkler, 2012b) have been devised in the last few decades to allow for a better evaluation of scientific literature, as "relative indicators relate the measure of scientometric elements with given units of a scientometric set to that of another set selected as absolute reference standard offering similar elements and units" (Vinkler, 2003, p. 691). These indicators – such as, for instance, the MECR, *Mean Expected Citation* Rate, or the MOCR, *Mean Observed Citation Rate* – are extensively described and tested in the last two papers cited above, as well as in Vinkler (2009) and in Glänzel et al. (2009). In general, the purpose of these indicators (defined "BMV indices" in Vinkler, 2012b) is to normalise the impact of journals (and possibly of single articles[3])

[2] Vanclay previously worked on the topic of JIF biases due to the two-year citation window (Vanclay, 2009).
[3] An example is the SNIP-Source Normalised Impact per Paper (Moed, 2010).

Table 1
Data on journals in the sample.

Title	IF 2010	Journal Country/Territory	Sum of published articles 1999–2010
Chemical Reviews	33.036	U.S.A.	1991
Angewandte Chemie-International Edition	12.730	Germany	16,631
Chemistry-A European Journal	5.476	Germany	9777
Langmuir	4.269	U.S.A.	21,305
Journal of Computational Chemistry	4.050	U.S.A.	2469
New Journal of Chemistry	2.631	U.K./France	3390
Bulletin of the Chemical Society of Japan	1.574	Japan	3426
Helvetica Chimica Acta	1.284	Switzerland	3259
Heteroatom Chemistry	1.044	U.S.A.	1145
Chinese Chemical Letters	0.775	Peoples' R. of China	4837
Research on Chemical Intermediates	0.715	Netherlands	995
Acta Chimica Sinica	0.611	P.R.C./Germany	4834
South African Journal of Chemistry	0.567	South Africa	283
Asian Journal of Chemistry	0.247	India	7154
Academy of Management Review	6.720	U.S.A.	810
Journal of Management	3.758	U.S.A.	550
Strategic Management Journal	3.583	U.S.A.	847
Research Policy	2.508	Netherlands	1304
Research-Technology Management	0.754	U.S.A.	1291
Canadian Journal of Administrative Sciences	0.714	U.S.A./Canada	361
International Journal of Technology Management	0.519	Switzerland	1107

Source: ISI-WoS-ThomsonReuters.

across scientific areas, fields and sectors; e.g. "normalising the citation(s) to a publication by their expected value is an acknowledged way to proceed" (Amez, 2012, p. 1460). This is due to the well-known fact that "[. . .] certain key statistics – such as [. . .] the mean citation ratio [. . .] – exhibit a large degree of variation across scientific fields" (Albarrán & Ruiz-Castillo, 2011, p. 48).

This brief overview clearly shows that means of citations are widely used as indices of performance of scientific literature; the variant adopted in this work has been studied to fit the use of the JIF made here.

This theoretical framework aims at bringing together current opinions and studies on the JIF, describing some critical issues related to its computation and uses, and also showing the need for further investigating its characteristics. Our purpose is to gather experimental evidence to support some of the points made above.

3. Methodology

The methodology adopted in this work is based on the study of a sample of journals over the 1999–2010 time period (relying on the in-house version of the ISI Journal Citation Reports® at the University of Turin). The sample has been chosen in order to be meaningful within the limits of an operable research effort. A number of journals from the ISI Journal Citation Reports® have been selected. The sample includes 14 journals from the ISI Category "Chemistry, multidisciplinary" (JCR Science Edition) and 7 journals from the ISI Category "Management" (JCR Social Sciences Edition), which represent just above 10% of the total population of journals in the two categories. Indeed, the ISI category "Chemistry, Multidisciplinary" included 147 journals in 2010 and 121 in 1999 (when it was labelled "Chemistry"), whereas "Management" comprised 144 journals in 2010 but only 61 in 1991; hence the choice of selecting 14 chemistry journals and 7 management journals to make up our dataset.

The two categories have been chosen since they are not extremely specific; therefore, they should be better suited to showing general results. The choice of journals has been made by performing a systematic sampling of the two categories, according to the following determinants:

- Presence within the ISI JCR throughout the analysed time period (1999–2010).
- Journal Impact Factors (according to ISI JCR 2010): journals have been chosen from the top of a category as well as from its lower end.
- Total number of articles published in the studied period: from highly productive journals to low production ones.
- Journal Country/Territory (according to ISI JCR data): origin differentiation, choosing journals from different geographical areas.
- Within the limits of these determinants, the choice has been performed randomly, with a certain amount of trial-and-error and excluding possible candidates not fitting the determinants (e.g. a journal might have had a suitable JIF but it might not have been present within the ISI JCR since 1999).

This rational system has been devised to ensure the selection of a representative sample.

Table 1 reports the titles of and some relevant data about the journals included in the sample.

Table 2
Descriptive statistics for JIF and MRC – data from 1999 to 2010.

Journal	JIF st. dev.	JIF mean	JIF CV	MRC st. dev.	MRC mean	MRC CV
Chemical Reviews	5.331	24.281	0.220	5.331	24.281	0.220
Angewandte Chemie-International Edition	1.537	9.593	0.160	1.144	6.783	0.169
Chemistry-A European Journal	0.435	4.900	0.089	0.503	4.028	0.125
Langmuir	0.489	3.539	0.138	0.487	3.293	0.148
Journal of Computational Chemistry	0.645	3.523	0.183	1.469	3.933	0.374
New Journal of Chemistry	0.310	2.598	0.119	0.337	1.947	0.173
Bulletin of the Chemical Society of Japan	0.189	1.512	0.125	0.200	1.137	0.176
Helvetica Chimica Acta	0.360	1.766	0.204	0.213	1.204	0.176
Heteroatom Chemistry	0.144	0.871	0.165	0.135	0.750	0.180
Chinese Chemical Letters	0.169	0.383	0.442	0.161	0.349	0.462
Research on Chemical Intermediates	0.158	0.644	0.246	0.158	0.521	0.303
Acta Chimica Sinica	0.178	0.656	0.272	0.113	0.435	0.259
South African Journal of Chemistry	0.141	0.397	0.354	0.132	0.288	0.458
Asian Journal of Chemistry	0.044	0.217	0.201	0.027	0.170	0.161
Academy of Management Review	1.401	4.762	0.294	1.374	4.817	0.285
Journal of Management	1.057	2.119	0.499	1.317	4.618	0.285
Strategic Management Journal	0.726	2.825	0.257	2.753	5.921	0.465
Research Policy	0.587	1.711	0.343	1.006	2.988	0.337
Research-Technology Management	0.160	0.497	0.322	0.094	0.233	0.404
Canadian Journal of Admin. Sciences	0.182	0.220	0.830	0.137	0.377	0.362
International J. of Technology Man.	0.124	0.302	0.411	0.103	0.406	0.254

Data mining was performed on Thomson Reuters – ISI Web of Knowledge® in May 2012. The obtained data include:

- Journal Impact Factors, in the time span ranging from 1999 to 2010, retrieved from the ISI-JCR.
- The Mean number of Received Citations per article per year (MRC) weighed for the age of the cited items, calculated on the basis of Web of Knowledge data. These values are obtained by calculating the total sum of citations received (by articles published in a given year) from publication up to 2011. Then, the total is divided by the number of articles multiplied by the "lifespan" of the articles (including their publication year) according to the formula:

$$\mathrm{MRC}_i = \frac{\sum_{y=i}^{2011} \mathrm{received\_citation}_y}{\mathrm{number\_of\_articles}_i \times (2012 - i)}$$

A further reason for the use of this MRC formula is to avoid age-dependence, clearly shown in e.g. Liang, Rousseau, and Zhong (2012), which would make the data unsuitable for the specific use made in the present work.

The MRCs are used here to measure the performance of journals in a specific year. The number of citations has been used in the past with the aim of measuring the performance of scientific products by several authors (see, for instance, Potì et al., 2012; Vieira and Gomes, 2010). Specific time spans have been used to eliminate problems regarding citation lags (see below).

Some descriptive statistics on the obtained datasets are presented in Table 2: standard deviation, mean and coefficient of variation for both the JIF and the MRC per article per year. Table 3, instead, contains the median of received citations for the year's articles.

Figs. 1 and 2 show the scatter plots of IF vs. MRC for Chemistry and Management journals, which complement the information provided by correlation coefficients.

The correlation coefficients are calculated from the two sets of data for different ranges and combinations, as described below. Pearson's Correlation Coefficients r as well as non-parametric coefficients (Spearman's Rank-Order Correlation Coefficients and Kendall's τ_α) are calculated. Although Kendall's τ_α might be considered redundant, it is taken into account in order to provide further confirmation of the obtained results.[4]

A first set of the three correlation coefficients is calculated for JIF vs. MRC for two time periods: 1999–2010 and 1999–2007. The second (shorter) time period has been chosen to eliminate the effects of the lag between articles being published and being cited, which exists in many cases. Thus, the last three years of the sample have been excluded from the calculation of the second group of coefficients.

A second set of correlation coefficients is calculated for the differences between years i and $i-1$ for what concerns JIF vs. MRC, ($\mathrm{JIF}_i - \mathrm{JIF}_{i-1}$ vs. $\mathrm{MRC}_i - \mathrm{MRC}_{i-1}$). Also in this case the calculations refer to two time periods: 2000–2010 and

[4] See, for instance, Franceschet (2010) for the use of Spearman's Rank-Order Coefficients with bibliometric indicators. Spearman's and Kendall's Coefficients have been calculated with the aid of Wessa (2012).

Table 3
Median of total received citations for the year's articles.

	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Chemical Reviews	210	185	133.5	170.5	117	141	119	70	87	65	46	26
Angewandte Chemie-International Edition	38	36	35	36	37	36	32	29	24	21	14	7
Chemistry-A European Journal	34	29	29	26	26	27	23	21	17	14	9	5
Langmuir	24	24	23	21	21	20	19	16	14	10	7	3
Journal of Computational Chemistry	16	16	16	17.5	13	18	12	8	9	6	5	2
New Journal of Chemistry	14	17	13	16	14	12	12	8	7	6	4	2
Bulletin of the Chemical Society of Japan	7	7	7	6	7	6	5	4	4	3	2	1
Helvetica Chimica Acta	11	12	12	10	8	8	8	5	4	3	2	1
Heteroatom Chemistry	4	6	5	6	4.5	5	4	3.5	3	2	1.5	0
Chinese Chemical Letters	1	1	1	2	2	1	1	1	2	1	1	0
Research on Chemical Intermediates	4	2.5	3	3	3	2	2	2	3	2	1	0
Acta Chimica Sinica	2	2	2	3	3	3	1	2	2	1.5	1	0
South African Journal of Chemistry	1	1.5	0.5	1	2	1	2	2	1	1	1	0
Asian Journal of Chemistry	1	1	1	1	1	1	0	0	0	0	0	0
Academy of Management Review	35	34	7	7	17	2.5	13	19	22.5	8	5.5	5.5
Journal of Management	43	41	43.5	25.5	34	24.5	21	20	25	9.5	5	3
Strategic Management Journal	60	67	64	42	41	42	35.5	23	22	10	6	2
Research Policy	19	21	16	23	18.5	17.5	17	18	10	6	4	1
Research-Technology Management	0	0	0	0	0	0	0	0	0	0	0	0
Canadian Journal of Admin. Sciences	1	1	3	1	1	1	0	0	1	1	1	0
International J. of Technology Man.	3	2	2	2	1	3	2	3	1	1	0	0

2000–2007. The aim of this second set of coefficients is to test the correlation existing between the yearly variations of the two magnitudes.

The third and last set of correlation coefficients is calculated for JIF_i vs. $(MRC_{i-1} + MRC_{i-2})/2$, i.e. the mean of the two values of MRC referring to the previous two years. This formula is more similar to the one used for the calculation of the JIF.

Table 4 reports the values of Pearson's Correlation Coefficients r calculated for all the journals, whereas Table 5 contains the values of Spearman's Correlation Coefficients and Table 6 those of Kendall's $\boldsymbol{\tau_\alpha}$. Values equal or above +0.500 are reported in bold to ensure better clarity.

The following section describes the results of our analysis.

4. Results

The calculated correlation coefficients display a wide range of values for all three sets of coefficients. Tables 4–6 also include the maximum and minimum values for each of the sets of coefficients (shown in italics at the bottom of the tables). The most consistent set of data is that of the third set of coefficients (JIF vs. average MRC for the previous two years) for Chemistry journals, which in most cases display values above +0.500.

In order to verify the significance of the obtained data, the correlation coefficients are then tested against a wide range of null hypotheses (listed in Table 7), from linear correlation to 0 correlation. As described above, the JIF and MRC are expected to be strongly correlated (in principle with linear correlation). The same is obviously expected to happen for what concerns their yearly variations. Following this rationale, the null hypotheses used to test the coefficients range from more stringent correlations (0.99 and 0.95) to weaker (0.66), weak (0.33) and no correlation (0).

The procedure adopted in this testing phase is as follows. Regarding Pearson's Correlation Coefficients, for both values of r and ρ_0 the transformation proposed by Fisher (1915, 1921) for small samples has been used. According to Fisher, r' is derived from r by using the formula:

$$r' = 0.5 \ln\left(\frac{1+r}{1-r}\right);$$

ρ'_0 is derived from ρ_0 by using the same formula; its values are reported in Table 7.

Both r' and ρ' are normally distributed, and have

$$\sigma = \sqrt{\frac{1}{n-3}};$$

A test statistic can thus be obtained with the usual formula:

$$Z = \frac{r' - \rho'_0}{\sigma}$$

(See also Howell, 2010, pp. 275 ff.)

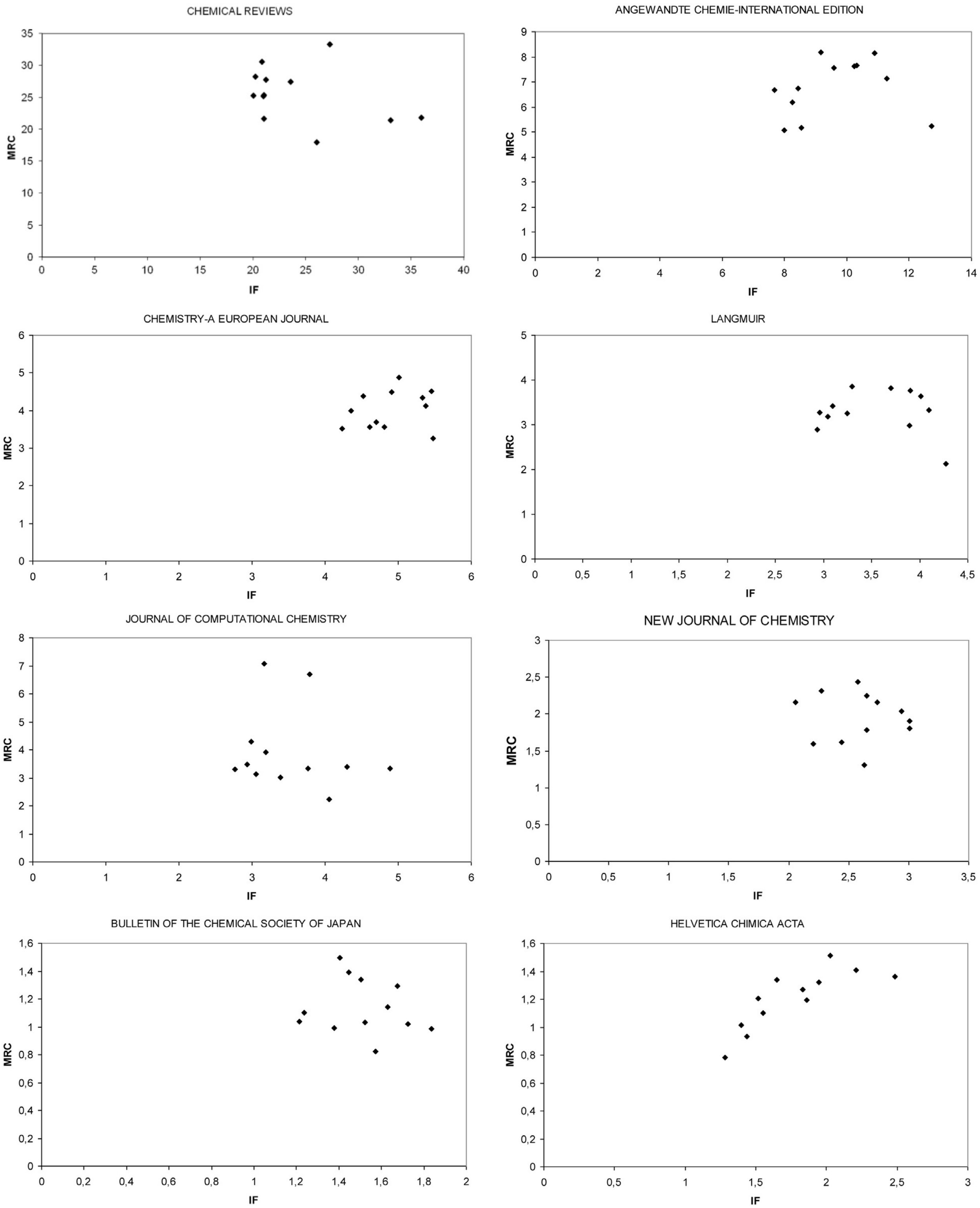

Fig. 1. Scatter plots of IF vs. MRC (Chemistry journals).

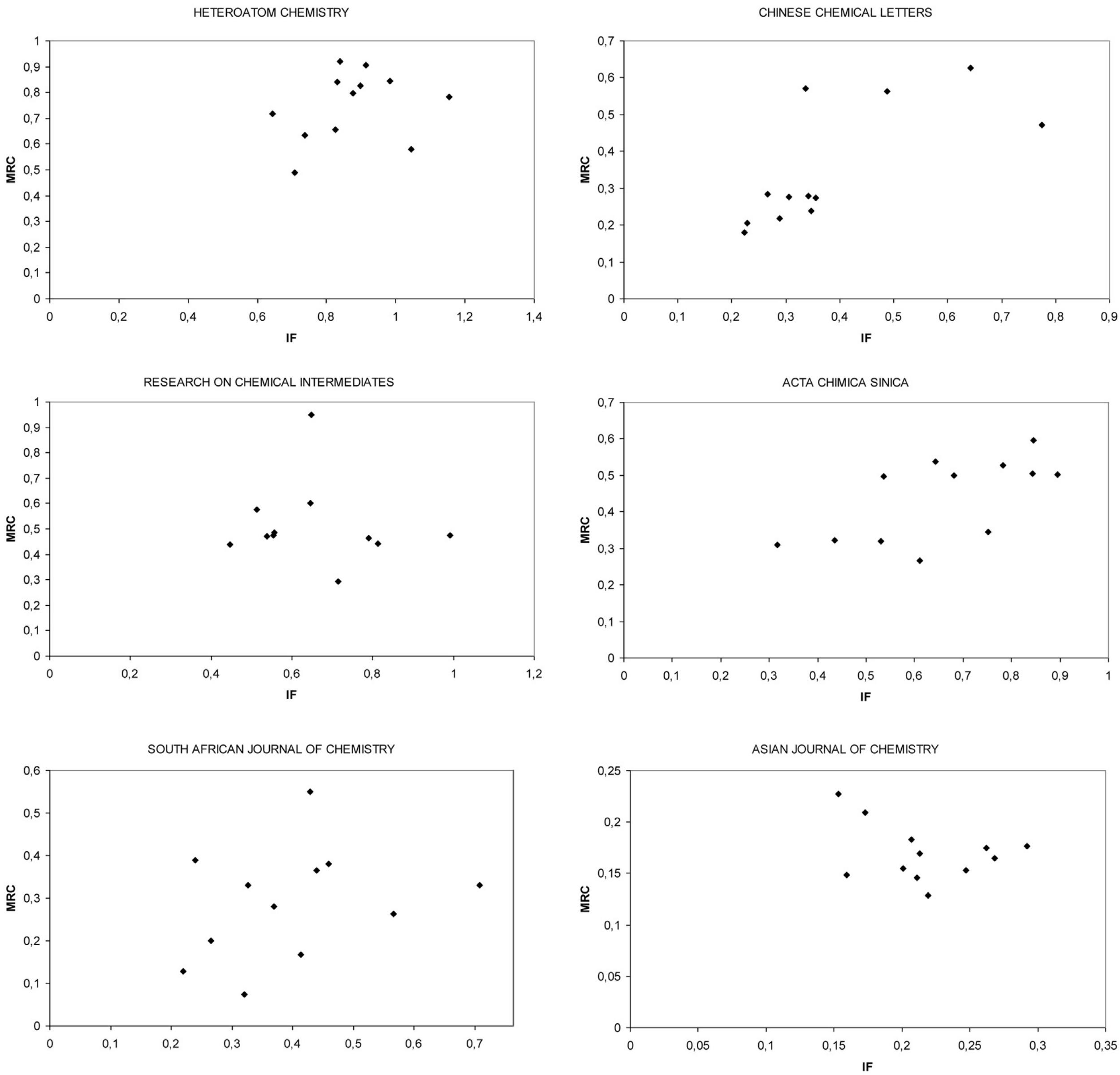

Fig. 1. (*Continued*).

For Spearman's Rank-Order Correlation Coefficients and Kendall's $\boldsymbol{\tau_\alpha}$, the same transformation has been adopted, following Fieller, Harley, and Pearson (1957), Fieller and Pearson (1961) and Pearson and Snow (1962). Based on the guidelines provided by these authors, standard deviations have been calculated as:

$$\sigma = \frac{1.03}{\sqrt{(n-3)}}$$

for Spearman's Rank-Order Correlation Coefficients and as:

$$\sigma = \frac{0.66}{\sqrt{(n-4)}}$$

for Kendall's $\boldsymbol{\tau_\alpha}$.

Tables 8–10 summarise the obtained results, showing the number of non-rejected correlation coefficients ($\alpha = 5\%$, two sided, critical value = ±1.96) for each null hypothesis and for each set of coefficients for the three calculated coefficients.

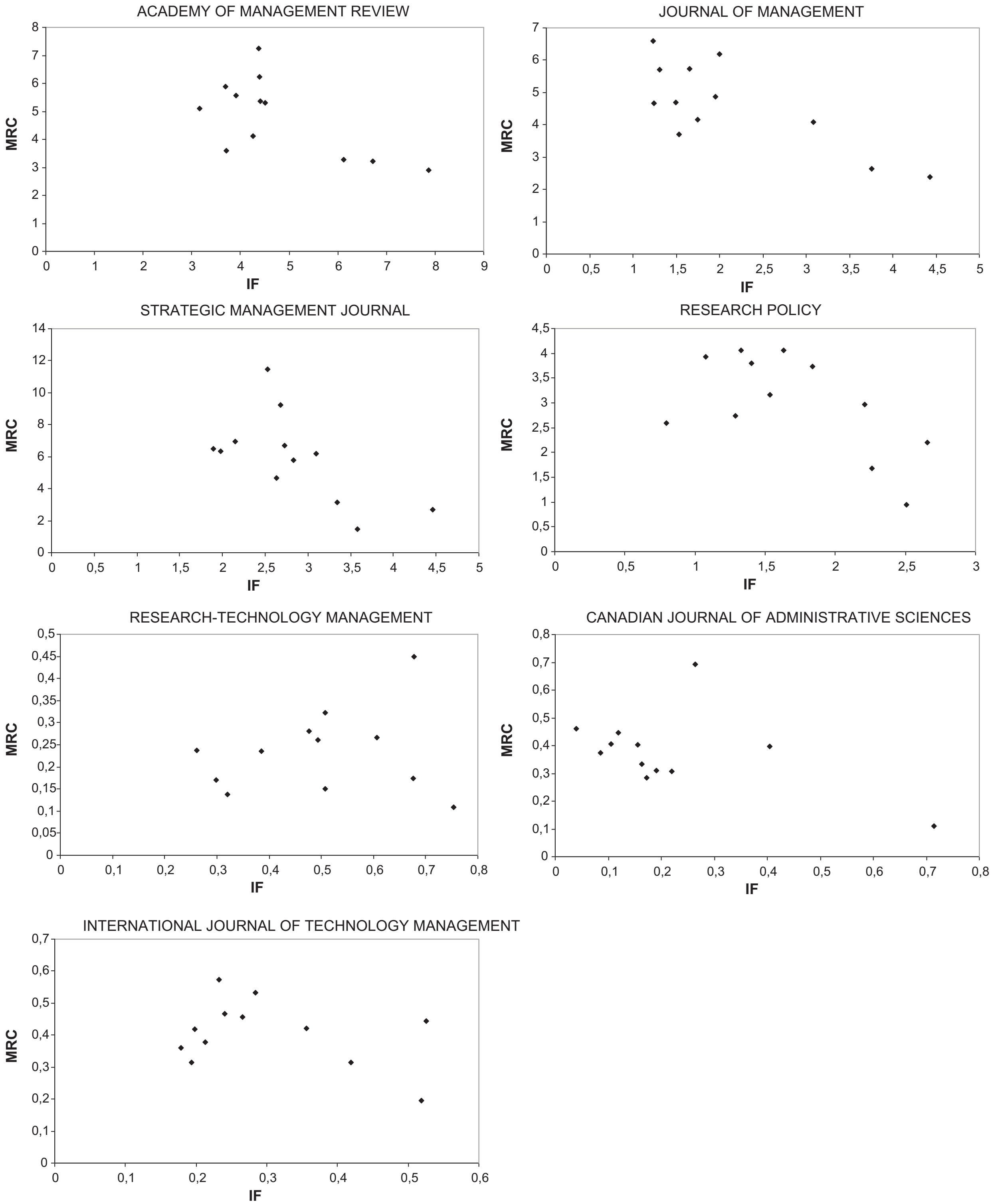

Fig. 2. Scatter plots of IF vs. MRC (Management journals).

Table 4
Pearson's correlation coefficients.

	JIF vs. MRC		$(JIF_i - JIF_{i-1})$ vs. $(MRC_i - MRC_{i-1})$		JIF vs. mean of MRC for 2 previous years	
	1999–2010	1999–2007	2000–2010	2000–2007	2001–2010	2001–2007
Chemical Reviews	−0.356	0.025	−0.272	−0.462	**0.512**	0.378
Angewandte Chemie-International Edition	0.187	**0.693**	−0.514	−0.170	**0.821**	**0.916**
Chemistry-A European Journal	0.233	**0.537**	0.032	−0.016	**0.922**	**0.956**
Langmuir	−0.152	**0.738**	−0.204	−0.242	**0.586**	**0.982**
Journal of Computational Chemistry	−0.155	−0.052	−0.398	−0.558	**0.688**	**0.923**
New Journal of Chemistry	−0.076	−0.039	0.065	−0.081	**0.637**	**0.826**
Bulletin of the Chemical Society of Japan	−0.159	−0.112	−0.196	−0.142	**0.870**	**0.787**
Helvetica Chimica Acta	**0.808**	**0.646**	−0.187	−0.238	**0.896**	**0.759**
Heteroatom Chemistry	0.280	**0.670**	−0.184	0.018	**0.633**	**0.661**
Chinese Chemical Letters	**0.720**	0.465	−0.066	0.407	**0.956**	0.034
Research on Chemical Intermediates	−0.142	0.003	0.309	**0.563**	0.365	−0.343
Acta Chimica Sinica	**0.655**	**0.813**	−0.018	−0.304	**0.897**	**0.948**
South African Journal of Chemistry	0.305	0.335	−0.218	−0.079	**0.698**	**0.642**
Asian Journal of Chemistry	−0.303	−0.265	−0.591	−0.666	**0.661**	**0.840**
Academy of Management Review	−0.638	0.311	−0.328	0.454	−0.234	−0.755
Journal of Management	−0.764	−0.077	−0.564	−0.245	−0.332	−0.490
Strategic Management Journal	−0.617	−0.006	−0.077	−0.103	−0.077	−0.103
Research Policy	−0.601	0.124	−0.123	−0.158	−0.123	−0.158
Research-Technology Management	0.164	**0.820**	0.379	**0.699**	0.050	0.405
Canadian Journal of Admin. Sciences	−0.481	0.320	−0.054	0.409	−0.304	0.039
International J. of Technology Man.	−0.353	0.403	0.071	−0.153	0.166	0.366
Maximum for Chemistry Journals	***0.808***	***0.813***	*0.309*	***0.563***	***0.956***	***0.982***
Minimum for Chemistry Journals	*−0.502*	*−0.265*	*−0.591*	*−0.666*	*−0.048*	*−0.343*
Maximum for Management Journals	*0.164*	***0.820***	*0.379*	***0.699***	*0.166*	*0.405*
Minimum for Management Journals	*−0.764*	*−0.077*	*−0.564*	*−0.245*	*−0.332*	*−0.755*

Bold: coefficients above 0.500; italics: maxima and minima for each set of coefficients.

Table 5
Spearman's Rank-Order Correlation Coefficients.

	JIF vs. MRC		$(JIF_i - JIF_{i-1})$ vs. $(MRC_i - MRC_{i-1})$		JIF vs. mean of MRC for 2 previous years	
	1999–2010	1999–2007	2000–2010	2000–2007	2001–2010	2001–2007
Chemical Reviews	−0.315	0.000	0009	−0.024	0.394	0.429
Angewandte Chemie-International Edition	0.413	**0.733**	−0.182	−0.024	**0.891**	**0.929**
Chemistry-A European Journal	0.245	**0.569**	0.196	0.024	**0.778**	**0.883**
Langmuir	0.126	**0.750**	−0.009	−0.192	0.491	**0.929**
Journal of Computational Chemistry	−0.133	0.133	−0.200	−0.262	0.224	**0.643**
New Journal of Chemistry	−0.049	0.050	−0.127	−0.190	**0.600**	**0.714**
Bulletin of the Chemical Society of Japan	−0.294	0.150	−0.218	−0.071	**0.891**	**0.893**
Helvetica Chimica Acta	**0.881**	**0.733**	−0.055	0.095	**0.915**	**0.786**
Heteroatom Chemistry	0.336	**0.750**	−0.227	0.143	**0.661**	**0.714**
Chinese Chemical Letters	**0.671**	0.433	0.032	0.443	**0.709**	0.143
Research on Chemical Intermediates	−0.172	0.243	0.337	**0.551**	−0.018	−0.286
Acta Chimica Sinica	**0.727**	0.750	0.145	−0.095	**0.867**	**0.679**
South African Journal of Chemistry	0.371	0.317	−0.328	−0.204	**0.661**	**0.750**
Asian Journal of Chemistry	−0.189	−0.367	−0.473	−0.429	**0.661**	**0.750**
Academy of Management Review	−0.476	0.250	−0.100	0.310	−0.370	−0.679
Journal of Management	−0.566	−0.050	−0.627	−0.429	−0.321	−0.500
Strategic Management Journal	−0.692	−0.283	−0.309	−0.214	−0.552	0.286
Research Policy	−0.497	0.150	−0.136	0.119	−0.030	**0.571**
Research-Technology Management	0.140	**0.800**	0.100	0.333	0.115	0.321
Canadian Journal of Admin. Sciences	−0.455	−0.267	−0.109	0.119	−0.091	0.214
International J. of Technology Man.	0.042	**0.683**	0.109	0.214	0.055	0.071
Maximum for Chemistry Journals	***0.881***	***0.750***	*0.337*	***0.551***	***0.915***	***0.929***
Minimum for Chemistry Journals	*−0.371*	*−0.367*	*−0.473*	*−0.429*	*−0.479*	*−0.643*
Maximum for Management Journals	*0.140*	***0.800***	*0.109*	*0.333*	*0.115*	***0.571***
Minimum for Management Journals	*−0.692*	*−0.283*	*−0.627*	*−0.429*	*−0.552*	*−0.679*

Bold: coefficients above 0.500; italics: maxima and minima for each set of coefficients.

Table 6
Kendall's τ_α.

	JIF vs. MRC		$(JIF_i - JIF_{i-1})$ vs. $(MRC_i - MRC_{i-1})$		JIF vs. mean of MRC for 2 previous years	
	1999–2010	1999–2007	2000–2010	2000–2007	2001–2010	2001–2007
Chemical Reviews	−0.182	0.000	0.055	0.000	0.378	0.429
Angewandte Chemie-International Edition	0.303	**0.556**	−0.127	0.000	**0.778**	**0.810**
Chemistry-A European Journal	0.198	0.423	0.073	0.036	**0.584**	**0.781**
Langmuir	0.061	**0.500**	0.037	−0.109	0.422	**0.810**
Journal of Computational Chemistry	−0.091	0.056	−0.164	−0.214	0.022	0.429
New Journal of Chemistry	−0.061	0.056	−0.091	−0.143	0.422	**0.524**
Bulletin of the Chemical Society of Japan	−0.242	−0.167	−0.091	0.000	**0.733**	**0.714**
Helvetica Chimica Acta	**0.697**	**0.500**	−0.055	0.071	**0.778**	**0.619**
Heteroatom Chemistry	0.242	**0.556**	−0.127	0.071	**0.511**	**0.619**
Chinese Chemical Letters	**0.515**	0.333	0.000	0.255	**0.556**	0.048
Research on Chemical Intermediates	−0.076	0.254	0.220	0.327	0.000	−0.238
Acta Chimica Sinica	**0.515**	**0.556**	0.091	−0.071	**0.689**	**0.524**
South African Journal of Chemistry	0.242	0.222	−0.257	−0.182	**0.511**	**0.619**
Asian Journal of Chemistry	−0.121	−0.278	−0.345	−0.357	**0.511**	**0.619**
Academy of Management Review	−0.394	0.111	−0.127	0.143	−0.289	−0.524
Journal of Management	−0.424	0.000	−0.527	−0.429	−0.244	−0.333
Strategic Management Journal	−0.515	−0.167	−0.091	−0.071	−0.422	0.143
Research Policy	−0.364	0.056	−0.164	0.000	−0.022	0.429
Research-Technology Management	0.121	**0.611**	0.091	0.214	0.111	0.238
Canadian Journal of Admin. Sciences	−0.394	−0.278	−0.091	0.071	−0.067	0.143
International J. of Technology Man.	0.015	0.444	0.164	0.286	0.022	0.048
Maximum for Chemistry Journals	***0.697***	***0.556***	*0.220*	*0.327*	***0.778***	***0.810***
Minimum for Chemistry Journals	*−0.242*	*−0.278*	*−0.345*	*−0.357*	*−0.333*	*−0.429*
Maximum for Management Journals	*0.121*	***0.611***	*0.164*	*0.286*	*0.111*	*0.429*
Minimum for Management Journals	*−0.515*	*−0.278*	*−0.527*	*−0.429*	*−0.422*	*−0.524*

Bold: coefficients above 0.500; italics: maxima and minima for each set of coefficients.

Table 7
List of null hypotheses.

Linear correlation	$H_0: \rho_0 = 0.99$[a]	$\rho' = 2.6467$
Stronger correlation	$H_0: \rho_0 = 0.95$	$\rho' = 1.8318$
Less strong correlation	$H_0: \rho_0 = 0.66$	$\rho' = 0.7928$
Weaker correlation	$H_0: \rho_0 = 0.33$	$\rho' = 0.3428$
Zero or no correlation	$H_0: \rho_0 = 0$	$\rho' = 0$

[a] Notice that due to the formula of Fischer's transformation $H_0: \rho_0 = 1$ cannot be used.

Table 8
Z test for Pearson's Correlation Coefficients vs. null hypotheses – number of not rejected coefficients significant at $\alpha = 5\%$, two sided, critical value = ±1.96.

	JIF vs. MRC		(JIFi - JIFi-1) vs. $(MRC_i - MRC_{i-1})$		JIF vs. Mean of MRC for previous 2 years	
	1999–2010	1999–2007	2000–2010	2000–2007	2001–2010	2001–2007
$H_0: \rho_0 = 0.99$ – Chemistry	0	0	0	0	0	3
$H_0: \rho_0 = 0.99$ – Management	0	0	0	0	0	0
$H_0: \rho_0 = 0.95$ – Chemistry	0	1	0	0	6	9
$H_0: \rho_0 = 0.95$ – Management	0	1	0	0	0	0
$H_0: \rho_0 = 0.66$ – Chemistry	7	10	1	6	12	10
$H_0: \rho_0 = 0.66$ – Management	1	6	1	3	1	5
$H_0: \rho_0 = 0.33$ – Chemistry	11	14	11	12	8	9
$H_0: \rho_0 = 0.33$ – Management	1	6	6	7	7	6
$H_0: \rho_0 = 0$ – Chemistry	10	10	14	14	3	5
$H_0: \rho_0 = 0$ – Management	3	6	7	7	7	6

Table 9
Z test for Spearman's Correlation Coefficients vs. null hypotheses – number of not rejected coefficients significant at $\alpha = 5\%$, two sided, critical value = ±1.96.

	JIF vs. MRC		$(JIF_i - JIF_{i-1})$ vs. $(MRC_i - MRC_{i-1})$		JIF vs. mean of MRC for previous 2 years	
	1999–2010	1999–2007	2000–2010	2000–2007	2001–2010	2001–2007
H_0: $\rho_0 = 0.99$ – Chemistry	0	0	0	0	0	2
H_0: $\rho_0 = 0.99$ – Management	0	0	0	0	0	0
H_0: $\rho_0 = 0.95$ – Chemistry	1	0	0	0	4	10
H_0: $\rho_0 = 0.95$ – Management	0	1	0	0	0	0
H_0: $\rho_0 = 0.66$ – Chemistry	8	13	3	11	12	13
H_0: $\rho_0 = 0.66$ – Management	1	4	2	5	2	5
H_0: $\rho_0 = 0.33$ – Chemistry	12	14	14	14	10	10
H_0: $\rho_0 = 0.33$ – Management	2	7	6	7	6	6
H_0: $\rho_0 = 0$ – Chemistry	12	9	14	14	5	9
H_0: $\rho_0 = 0$ – Management	6	6	6	7	7	7

Table 10
Z test for Kendall's τ_α vs. null hypotheses – number of not rejected Coefficients Significant at $\alpha = 5\%$, two sided, critical value = ±1.96.

	JIF vs. MRC		$(JIF_i - JIF_{i-1})$ vs. $(MRC_i - MRC_{i-1})$		JIF vs. mean of MRC for previous 2 years	
	1999–2010	1999–2007	2000–2010	2000–2007	2001–2010	2001–2007
H_0: $\rho_0 = 0.99$ – Chemistry	0	0	0	0	0	0
H_0: $\rho_0 = 0.99$ – Management	0	0	0	0	0	0
H_0: $\rho_0 = 0.95$ – Chemistry	0	0	0	0	0	0
H_0: $\rho_0 = 0.95$ – Management	0	0	0	0	0	2
H_0: $\rho_0 = 0.66$ – Chemistry	3	9	0	2	12	13
H_0: $\rho_0 = 0.66$ – Management	0	2	0	2	0	5
H_0: $\rho_0 = 0.33$ – Chemistry	10	13	11	13	11	12
H_0: $\rho_0 = 0.33$ – Management	2	6	5	6	4	6
H_0: $\rho_0 = 0$ – Chemistry	11	11	14	14	5	10
H_0: $\rho_0 = 0$ – Management	6	6	6	7	7	7

A quick overview of the results shows that: H_0: $\rho_0 = 0.99$ and 0.95 are almost always rejected for the three calculated coefficients (with some exceptions for the third set of coefficients for Chemistry journals); H_0: $\rho_0 = 0.66$ is not rejected in many cases; H_0: $\rho_0 = 0.33$ and 0 are not rejected in most or all cases.

5. Discussion and conclusions

The aim of this work is to measure the correlation between the JIF (Journal Impact Factor) and the MRC (Mean Received Citations) in their evolution through time during a given set of years. This has been done by analysing a sample of journals, considering "citation performance" as a proxy for the average quality of the journals' yearly content, and then testing its significance, in order to contribute to the debate on how the JIF is used in the evaluation of scientific literature.

This paper does not claim to be exhaustive in its conclusions; nevertheless, some interesting observations can be made. The correlation coefficients for JIF vs. MRC are generally not very high, and in most cases far from a condition of linearity; this is notable in particular for Spearman's coefficients and Kendall's τ_α, which are less affected by the skewed distribution. The results are even more striking if we consider the correlations between yearly variations of the JIF and MRC. In this case, there is no coefficient above +0.500 for Kendall's τ_α, only one for Spearman's coefficients and only two for Pearson's coefficients (in the 2000–2007 time span).

The last set of coefficients (JIF vs. average MRC of the previous two years) instead displays higher coefficients for Chemistry journals than the former two sets. This result is interesting as the formula used for citations closely resembles that used for the calculation of the JIF. Thus, the evidence indicates that the JIF *for Chemistry journals* measures more closely the performance of the two years prior to its release. Conversely, Management journals do not present the same pattern for what concerns this last set, with most coefficients around or below 0.

The analysis of scatter plots for JIF vs. MRC (Figs. 1 and 2) shows that, for Chemistry journals, the distribution is almost always far from linear. One case (Helvetica Chimica Acta) displays a trend which is closer to linearity. This journal also shows higher correlation coefficients for the first set (interestingly not for the second-correlation between yearly changes). Nevertheless, it is difficult to infer a linear correlation from the scatter plots. When looking at Management journals, the scatter plot evidence is even more remarkable, as in most cases the plots exhibit a pattern closer to inverse linear correlation: most coefficients for JIF vs. MRC in the 1999–2010 period are in fact negative.

The results deriving from the testing of the obtained coefficients against null hypotheses are similar in all occurrences, despite the different nature of the calculated coefficients (parametric vs. non-parametric).

For the first two series (those calculated for the correlation between the two time series and for yearly changes of the two magnitudes[5]) H_0: $\rho_0 = 0.99$ is always rejected for the three coefficients and H_0: $\rho_0 = 0.95$ is always rejected but in two cases for Pearson's coefficients and in another two cases for Spearman's coefficients (at 5% significance). The tables show the results for H_0: $\rho_0 = 0.66$, 0.33 and 0.

The difference existing between the two groups of journals (Chemistry journals and Management journals) is noteworthy, and the same is true also for the two time periods under investigation (1999/2000/2001–2007 and 1999/2000/2001–2010). The correlation coefficients not considering the last three years included in the examined database show that the null hypothesis is not rejected – for all three coefficients – in a wider number of cases in comparison to the other set of coefficients (i.e. those referring to the longer time span). This indicates that the data might be biased due to the time span that often occurs between publication and being cited.

Management journals display a much lower share of non-rejected coefficients than Chemistry journals – again for all three calculated coefficients – in particular for H_0: $\rho_0 = 0.66$. This aspect is even more interesting if we bear in mind that the JIF considers citations from the two most recent years.

The correlation coefficients of the third group – those calculated for JIF_i vs. $(MRC_{i-1} + MRC_{i-2})/2$ – are not rejected in a few cases regarding Chemistry journals also for H_0: $\rho_0 = 0.99$ and H_0: $\rho_0 = 0.95$, at least for what concerns Pearson's and Spearman's coefficients.

As described above, these features are not completely surprising, considering that the formula closely resembles the one used to calculate the JIF for year *i*. These results might add further evidence to support the opinion expressed by Ingwersen (2012), who suggests using a *diachronic* JIF, rather than its current formulation. Nevertheless, this set of correlation coefficients is characterised by the most noticeable differences between the two groups of journals. H_0: $\rho_0 = 0.99$ and H_0: $\rho_0 = 0.95$ are always rejected for both time spans for Management Journals (with the exception of Kendall's $\boldsymbol{\tau_\alpha}$ in two cases); H_0: $\rho_0 = 0.33$ is rejected in one case for Pearson's coefficients, in two cases for Spearman's coefficients (one in each time span) and in some cases for Kendall's $\boldsymbol{\tau_\alpha}$ for what concerns Management journals, while it is not rejected in most cases for Chemistry journals. Interestingly H_0: $\rho_0 = 0$ is not rejected in fewer cases than for the previous null hypotheses.

The results of the experimental work carried out in the present study answer the initial research question, providing evidence to support the idea that the correlation between the evolution of the JIF and MRC from year to year is not strongly linear. Furthermore, it emerges that it is dangerous to infer the value of journal articles from a specific year using only the related JIF.

Instead, the JIF seems more closely linked – at least in some cases – to the Citedness of the previous two years. Chemistry and Management journals display great differences, a further clue suggesting that the JIF should not be used as a universal indicator of scientific production quality, at least for single scientific products. More generally, the findings described above confirm the importance of avoiding misuses of the JIF. In particular, the JIF for a given year should not be directly attached to the articles published in that year as a quality measure of single articles. The JIF was devised for a very specific purpose, as Garfield's statements reported above clearly indicate. The results presented in this article contribute to supporting the opinion that one should not go further in its use.

Acknowledgements

The present work has benefited from suggestions on methodology from Prof. Laura Magazzini (University of Verona), who is also acknowledged for her scientific collaboration, and from comments by two anonymous referees, who are acknowledged as well. Further acknowledgements go to: my advisor Prof. Salvatore Coluccia and Prof. Livio Battezzati (University of Torino) for supporting this research field in the recent past; Prof. Secondo Rolfo (CNR-CERIS) for supporting this research field, as well as for his scientific collaboration; Dr. M. Coccia and the colleagues at CNR-CERIS for their scientific collaboration. Usual disclaimer applies.

References

Albarrán, P., & Ruiz-Castillo, J. (2011). References made and citations received by scientific articles. *Journal of the American Society for Information Science and Technology*, *62*(1), 40–49.

Albarrán, P., Crespo, J. A., Ortuño, I., & Ruiz-Castillo, J. (2010). A comparison of the scientific performance of the U.S. and the European Union at the turn of the 21st century. *Scientometrics*, *85*(1), 329–344.

Amez, L. (2012). Citation measures at the micro level: Influence of publication age, field, and uncitedness. *Journal of the American Society for Information Science and Technology*, *63*(7), 1459–1465.

Archambault, E., & Larivière, V. (2009). History of the journal impact factor: Contingencies and consequences. *Scientometrics*, *79*(3), 635–649.

Balaban, A. T. (2012). Positive and negative aspects of citation indices and journal impact factor. *Scientometrics*, *92*(2), 241–247.

Bensman, S. J. (2007). Garfield and the impact factor. *Annual Review of Information Science and Technology*, *41*(1), 93–155.

Bensman, S.J. (2007). Garfield and the Impact Factor: The creation, utilization, and validation of a citation measure: Part 2, The probabilistic, statistical, and sociological bases of the measure, unpublished, available at: http://garfield.library.upenn.edu/bensman/bensmanegif22007.pdf (accessed December 2012).

Bensman, S. J. (2012). The impact factor: Its place in Garfield's thought, in science evaluation, and in library collection management. *Scientometrics*, *92*(2), 263–275. http://dx.doi.org/10.1007/s11192-011-0601-9

[5] That is to say, absolute values of the two sets of data along a time series, and changes of the same sets between year *i* − 1 and *i*.

Campanario, J. M. (2012). Some research ideas on Journal Impact Factor as a crucial topic in science dynamics. *Scientometrics*, *92*(2), 293–295. http://dx.doi.org/10.2007/s11192-012-0650-8

Fieller, E. C., & Pearson, E. S. (1961). Tests for rank correlation coefficients. II. *Biometrika*, *48*(1/2), 29–40.

Fieller, E. C., Harley, H. O., & Pearson, E. S. (1957). Tests for rank correlation coefficients. I. *Biometrika*, *44*(3/4), 470–481.

Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*, *10*(4), 507–521.

Fisher, R. A. (1921). On the "Probable Error" of a coefficient of correlation deduced from a small sample. *Metron*, *I*(4), 3–32.

Franceschet, M. (2010). Journal influence factors. *Journal of Informetrics*, *4*(3), 239–248.

Garfield, E. (1955). Citation indexes for science: A new dimension in documentation through association of ideas. *Science*, *122*(3159), 108–111.

Garfield, E. (1972). Citation analysis as a tool for journal evaluation. *Science*, *178*, 471–479.

Garfield, E. http://thomsonreuters.com/products_services/science/free/essays/impact_factor/ (accessed 20.07.12)

Garfield, E. (1997). The impact factor and using it correctly. *Der Unfallchirurg*, *48*(2), 413.

Garfield, E. (2005). The agony and the ecstasy – The history and meaning of the journal impact factor. In *Paper presented by Eugene Garfield at international congress on peer review and biomedical publication* Chicago, September 16.

Garfield, E. (2006). The history and meaning of the journal impact factor. *Journal of the American Medical Association*, *295*(1), 90–93.

Garfield, E., & Sher, I. H. (1963). New factors in the evaluation of scientific literature through citation indexing. *American Documentation*, *14*(3), 195–201.

Garfield, E., & Sher, I. H. (1966). New tools for improving and evaluating the effectiveness of research. In M. C. Yovits, D. M. Gilford, R. H. Wilcox, E. Staveley, & H. D. Lemer (Eds.), *Research program effectiveness, Proceedings of the conference sponsored by the office of naval research* Washington, DC, July 27–29, 1965, New York, Gordon and Breach, (pp. 135–146).

Glänzel, W., Thijs, B., Schubert, A., & Debackere, K. (2009). Subfield-specific normalized relative indicators and a new generation of relational charts: Methodological foundations illustrated on the assessment of institutional research performance. *Scientometrics*, *78*(1), 165–188.

Herranz, N., & Ruiz-Castillo, J. (2012). Sub-field normalization in the multiplicative case: Average-based citation indicators. *Journal of Informetrics*, *6*(XX), 543–556.

Howell, D. C. (2010). *Statistical methods for psychology*. Belmont, USA: Cengage Wadsworth. ISBN 978-0-495-59784-1.

Ingwersen, P. (2012). The pragmatics of a diachronic journal impact factor. *Scientometrics*, *92*(2), 319–324. http://dx.doi.org/10.1007/s11192-012-0701-1

Leydesdorff, L. (2006). Can scientific journals be classified in terms of aggregated journal–journal citation relations using the journal citation reports? *Journal of the American Society for Information Science and Technology*, *57*(5), 601–6013.

Leydesdorff, L. (2008). *Caveats* for the use of citation indicators in research and journal evaluations. *Journal of the American Society for Information Science and Technology*, *59*(2), 278–287.

Liang, L., Rousseau, R., Zhong, Z. (2012), Non-English journals and papers in physics and chemistry: Bias in citations? Scientometrics, in press

Moed, H. F. (2010). Measuring contextual citation impact of scientific journals. *Journal of Informetrics*, *4*(3), 265–277.

Moed, H. F., Colledge, L., Reedijk, J., Moya-Anegon, F., Guerrero-Bote, V., Plume, A., et al. (2012). Citation-based metrics are appropriate tools in journal assessment provided that they are accurate and used in an informed way. *Scientometrics*, *92*(2), 367–376. http://dx.doi.org/10.1007/s11192-012-0679-8

Mutz, R., & Daniel, H.-D. (2012). Skewed citation distribution and bias factor: Solutions to two core problems with the journal impact factor. *Journal of Informetrics*, *6*(2), 169–176.

Pearson, E. S., & Snow, B. A. S. (1962). Tests for rank correlation coefficients. III. Distribution of the Transformed Kendall coefficient. *Biometrika*, *49*(1/2), 185–191.

Pendlebury, D. A. (2009). The use and misuse of journal metrics and other citation indicators. *Archivum Immunologiae et Therapiae Experimentalis*, *57*(1), 1–11.

Pendlebury, D. A., & Adams, J. (2012). Comments on a critique of the Thomson Reuters journal impact factor. *Scientometrics*, *92*(2), 395–401.

Potì, B., Cerulli, G., & Romagnosi, S. (2012). The different impact on scientific output of academic-owned vs. academic-invented patents: Evidence from a panel of Italian academic inventors. *International Journal of Business Innovation and Research*, *64*, 433–457.

Rousseau, R. (2012). Updating the journal impact factor or total overhaul? *Scientometrics*, *92*(2), 413–417. http://dx.doi.org/10.1007/s11192-012-0649-1

Seglen, P. O. (1994). Causal relationship between article citedness and journal impact. *Journal of the American Society for Information Science and Technology*, *45*(1), 1–11.

Seglen, P. O. (1997). Why the impact factor of journals should not be used for evaluating research. *British Medical Journal*, *314*(7079), 498–502.

The PLoS Medicine Editors. (2006). The impact factor game. *PLoS Medicine*, *3*(6), pp. e291 (0707–0708).

van Leeuwen, T. (2012). Discussing some basic critique on Journal Impact Factors: Revision of earlier comments. *Scientometrics*, *92*(2), 443–455.

Vanclay, J. K. (2009). Bias in the journal impact factor. *Scientometrics*, *78*(1), 3–12.

Vanclay, J. K. (2011). Impact factor: Outdated artefact or stepping-stone to journal certification? *Scientometrics*, *92*(2), 211–238. http://dx.doi.org/10.1007/s11192-011-0561-0

Vieira, E. S., & Gomes, J. A. N. F. (2010). Citations to scientific articles: Its distribution and dependence on the article features. *Journal of Informetrics*, *4*(1), 1–13.

Vinkler, P. (2003). Relations of relative scientometric indicators. *Scientometrics*, *58*(3), 687–694.

Vinkler, P. (2009). Introducing the Current Contribution Index for characterizing the recent, relevant impact of journals. *Scientometrics*, *79*(2), 409–420.

Vinkler, P. (2012a). The Garfield impact factor, one of the fundamental indicators in scientometrics. *Scientometrics*, *92*(2), 471–483. http://dx.doi.org/10.1007/s11192-012-0688-7

Vinkler, P. (2012b). The case of scientometricians with the "absolute relative" impact indicator. *Journal of Informetrics*, *6*(2), 254–264.

Wessa, P. (2012). Free statistics software, office for research development and education, version 1.1.23-r7, URL http://www.wessa.net/

Wilhite, A. W., & Fong, E. A. (2012). Coercive citation in academic publishing. *Science*, *335*(6068), 542–543.

Wu, J. (2013). Investigating the universal distribution of normalized indicators and developing field-independent index. *Journal of Informetrics*, *7*(1), 63–71.