**Noname manuscript No.**
(will be inserted by the editor)

# Data-driven Journal Meta-Ranking in Business and Management

**Gianfranco Ennas · Battista Biggio ·
Maria Chiara Di Guardo**

Received: date / Accepted: date

**Abstract** Ranking journals is a longstanding problem and can be addressed quantitatively, qualitatively or using a combination of both approaches. In the last decades, the Impact Factor (*i.e.*, the most known quantitative approach) has been widely questioned, and other indices have thus been developed and become popular. Previous studies have reported strengths and weaknesses of each index, and devised meta-indices to rank journals in a certain field of study. However, the proposed meta-indices exhibit some intrinsic limitations: (i) the indices to be combined are not always chosen according to well-grounded principles; (ii) combination methods are usually unweighted; and (iii) some of the proposed meta-indices are parametric, which requires assuming a specific underlying data distribution. We propose a data-driven methodology that linearly combines an arbitrary number of indices to produce an aggregated ranking, using different techniques from statistics and machine learning to estimate the combining weights. We additionally consider correlations between indices and meta-indices, to quantitatively evaluate their differences. Finally, we empirically show that the considered meta-indices are also robust to significant perturbations of the values of the combined indices.

**Keywords** journal quality evaluation · combining journal ranking indices · meta-indices

Gianfranco Ennas
Department of Business and Economics, Cagliari, Italy
Tel.: +39 0706753450
E-mail: gianfrancoennas@unica.it

Battista Biggio
Department of Electrical and Electronic Engineering, Cagliari, Italy

Maria Chiara Di Guardo
Department of Business and Economics, Cagliari, Italy

## 1 Introduction

Ranking academic journals is an issue that affects many players, especially in academia; *e.g.*, scholars choosing among potential outlets for their research, departments to measure their productivity and to ensure funding success. It also affects the non-academic world, such as publishers aiming to evaluate the quality of their journals, professional societies, practitioners and funding organizations (Vokurka, 1996; Falagas et al, 2008). Obtaining a reliable journal ranking is a longstanding problem and its intrinsic difficulty relies on the multi-dimensionality of the quality concept. In fact, research quality can be measured qualitatively, quantitatively, or using a combination of both approaches (referred to as *hybrid approach* in the following). Qualitative approaches consist of ranking journals according to their perceived quality and reputation, *e.g.*, by interviewing a qualified sample of experts to rate journals in a particular field of study (Mylonopoulos and Theoharakis, 2001; Peffers and Ya, 2003; Sellers et al, 2004). Quantitative methods, instead, provide indices depending on the number of published articles in a journal, and the corresponding number of citations (Seglen, 1997; DuBois and Reeb, 2000; Hodge and Lacasse, 2010); *i.e.*, they evaluate two main aspects called respectively *size* and *impact* (Leydesdorff, 2009). Despite these methods capture only some aspects of the quality concept, they are easy to compute and are thus used as proxy measures to evaluate a journal quality. However, it should be kept in mind that quantitative aspects are not necessarily connected to qualitative ones, *e.g.*, publishing an article in a journal that publishes many articles does not make it automatically a high-quality article. To overcome this limitation, hybrid approaches have been also proposed that combine experts' opinions and quantitative approaches to better capture both aspects.

Since its proposal, the Impact Factor (IF) has been widely used as a quantitative approach; however, several limitations related to the (mis)use of this index have emerged (Sect. 2.1) (Seglen, 1997; Lancho-Barrantes et al, 2010; Bornmann et al, 2012). For this reason, alternative indices to the IF, such as the SCImago Journal Ranking (SJR), the H-index, the Eigenfactor Score, the Integrated Impact Indicator (I3) and the Source Normalized Impact per Paper (SNIP) have been developed and become popular in academia, although they exhibit other kinds of limitations (Sect. 2.2) (Leydesdorff, 2009; Mingers and Leydesdorff, 2015). This has favored the development of meta-indices that can mitigate issues specific to the use of each base index (Vanclay, 2008; Hodge and Lacasse, 2010; Bador and Lafouge, 2010; Theu$\beta$l et al, 2014; Tsai, 2014). In particular, three main issues have emerged: (i) indices to be combined are not always chosen according to well-defined guidelines, affecting the reliability of the corresponding meta-index; (ii) combination methods are usually unweighted; and (iii) some of the proposed meta-indices are parametric, requiring specific assumptions on the underlying data distribution (Sect. 2.3).

The contribution of our work is threefold: (i) to overcome the aforementioned limitations, we propose a data-driven methodology that estimates a weighted, linear combination of an arbitrary number of indices (potentially also including experts' opinions), yielding a more principled, aggregated journal ranking; (ii) we measure correlations between indices and meta-indices to compare their relative performance; and (iii) we provide an empirical analysis of the considered meta-indices against random perturbations of the values of the combined indices, to assess the stability of the corresponding journal rankings (Sect. 3). We empirically validate our approach on two journal databases obtained by combining journals indexed in Business and Management from the Thomson Reuters Journal Citation Report (JCR) and from SCImago Journal & Country Rank (SJR) (Sect. 4). We finally discuss conclusions and future research directions in Sect. 5.

## 2 Background

In this section, we start analyzing the main limitations of the Impact Factor and discuss other citation indices, including the SJR and the H-Index. We then highlight the need of combining indices and review the main meta-indices proposed so far.

### 2.1 Limitations of the Impact Factor

As stated in the introduction, the various indices used to measure research quality have limitations in nature or in capturing the different dimensions of the quality concept. In particular, the Impact Factor (IF) index, was developed by Eugene Garfield (Garfield, 2006) now is a product of Thomson Reuters Corporation and by definition is: *"the average number of times articles from the journal published in the past two years have been cited in the JCR year"*.

$$\mathrm{IF} = \frac{\text{citations in } Y_n \text{ of documents published in } Y_{n-1} + Y_{n-2}}{\text{citable items in } Y_{n-1} + Y_{n-2}} \quad . \qquad (1)$$

It has been criticized for multiple reasons such: (i) small research fields tend to lack journals with high impact; (ii) citation rates of articles determine journal impact but not vice versa; (iii) IF is a function of the number of references per article in the research field; (iv) in some journals (*e.g.* Nature) letters and correspondence are considered citations and, of course, they inflate the index; (v) journal impact factors are not statistically representative of individual journal articles. In other words, different scientific areas, fields and micro-fields of study have different citation habits; (vi) IF have a limited coverage, in particular in the Social Sciences and Arts and Humanities (Seglen, 1997; Lancho-Barrantes et al, 2010; Bornmann et al, 2012). Even, the academic journal Scientometrics dedicated a special issue in the 2012 aimed to evaluate IF problems also in comparison to its counterparts. Despite all these limitations, IF is widely used, for three reasons: the first is a *path dependence*, as since its introduction scholars and editors have learnt to use it, and without any strong alternative it has become a "*de facto* standard"; the second reason is that the computation method of this index is intrinsically easy to understand; and the third reason is that, due to its simplicity and growing popularity, it has been also (mis)used to measure the overall impact of journals. However, it is worth reminding that the IF was originally thought to a different end, *i.e.*, to choose journals to be included in the newly created science citation index by Garfield and Sher (1963).

### 2.2 A Brief Review of Citation Indices

During the last years, several indices have become popular, especially after that Elsevier launched the Scopus database as an alternative to the Thomson Reuters Web-of-Science (WoS) databases (Vieira and Gomes, 2009). The Scopus database includes more titles, from more countries, and the considered journals are published in a greater variety of languages than the Thomson Reuters ones (Leydesdorff et al, 2010). In this work, we consider a representative set of the most known and used citation indices, including: the Eigenfactor Score, the Article Influence Score (AIS), the SCImago Journal & Country Rank index (SJR), the H-Index, the Immediacy Index (Imm), the Impact Per Publication (IPP) and the Source Normalized Impact Per Publication (SNIP).

**Eigenfactor Score.** It is implemented in the WoS database. It rates journals according to the number of incoming citations, giving more emphasis to citations coming from highly-ranked journals, rather than equally weighting them (Bergstrom, 2007). The main advantages of this index include the fact that it excludes journal self-citations, it is freely available, it compensates for citation differences across disciplines, and it uses the structure of the entire network to evaluate the importance of each journal. (Mingers and Leydesdorff, 2015; Eigenfactor, 2015).

**Article Influence Score.** It is obtained by dividing the Eigenfactor score by the fraction of articles recorded in JCR and published in a specific journal over the last five years. Article Influence scores are normalized so that the mean article in the entire Thomson JCR database has an article influence of 1.00 (Mingers and Leydesdorff, 2015).

**SJR.** It is based on a similar idea to that behind the Eigenfactor score. In particular, it is calculated in the Scopus database as the average number of weighted citations received in the selected year by documents that have been published in the selected journal during the previous three years (SJR, 2007). Conceptually, this index is easy to understand, despite its calculation is rather complicated; in fact, it relies on an iterative Page Rank algorithm that distributes prestige values among the journals until a steady-state solution is reached. The SJR algorithm starts giving an identical amount of prestige to each journal, then this prestige is redistributed in a process where journals transfer their achieved prestige to each other through citations. The process ends up when the difference between journal prestige values in consecutive iterations does not exceed a minimum threshold value any more. Hence, SJR does not only take into account the number of citations, but it also provides a higher score to journals having articles cited by the most prestigious journals. Accordingly, the SJR index essentially reduces influence of self-citations (in particular, they can not exceed 33% of the total number of citations), *i.e.*, prestige can be transferred to a journal only by the other journals, and not self-transferred (González-Pereira et al, 2010). Differently from the IF, the SJR is normalized by considering all articles from a journal (and not only the citable ones). Recently, it has been shown that a correlation between SJR and IF exists; in particular, Falagas et al (2008) have shown that half of the journals in the IF top 100 list are placed within a reasonable range of ranking places in the SJR indicator journals list. Furthermore, IF and SJR are directly comparable as they exploit a similar time window, respectively, of 2 and 3 years. Lastly, it has been claimed that SJR normalizes for size a bit more strongly than the IF, and it is highly correlated with experts' opinions (Hodge and Lacasse, 2010).

**H-Index.** It was proposed about a decade ago by Hirsch and rapidly gained a widespread acceptance. Its initial aim was to measure scholars' productivity and citation impact: "a scientist has index $h$ if $h$ of his or her $Np$ papers have at least $h$ citations each and the other $(Np - h)$ papers have less than $h$ citations each" (Hirsch, 2005). Subsequently Braun et al (2006) have shown how that this index could be successfully applied also to journals, rather than only to scientists. The main advantage of the H-index is that, conversely to other indices (such as the SJR, the IF, and the Eigenfactor Score), it does not have an arbitrarily-fixed time horizon. It can be theoretically calculated since the creation of the journal, even if this may not be appropriate. This index is also insensitive to an excess of non-cited/highly-cited articles, as it does not rely upon the computation of any mean value. Note however that this characteristic could be seen also as a disadvantage (Braun et al, 2006; Leydesdorff, 2009).

**Immediacy Index.** According to the Institute for Scientific Information (ISI) Web of Knowledge, this index corresponds to the average number of citations received by an article (published in a given journal) during its first year of publication. By construction, this index evaluates the potential, immediate impact of a paper published in a given journal. Notably, frequently-issued journals could be advantaged, as articles that are published earlier during the year have a higher probability of being cited than articles published later in the year. The Immediacy Index is published annually in the Journal Citation Reports.

**Impact Per Publication (IPP).** It measures the ratio of citations in a year to scholarly papers published in the three previous years, divided by the number of scholarly papers published in those same years. Taking into account the same peer-reviewed scholarly papers only in both the numerator and denominator of the equation provides a fair impact measurement of the journal, while reducing the chance of manipulation (Scopus, 2015).

**Source Normalized Impact Per Publication (SNIP).** It is computed by normalizing the IPP on the number of citations in the corresponding subject field. The SNIP thus measures contextual citation impact by weighting citations based on the total number of citations in a subject field. It can be defined as the ratio of a journal's citation count per paper and the citation potential in its subject field (Scopus, 2015). Generally, the main advantage of this index is that the reference set of journals is defined at the time specifically for the collection of papers being evaluated (Mingers and Leydesdorff, 2015). Thus, the impact of a single citation receives a higher value in low-cited research areas and *vice versa*, making the SNIP more reliable than the IF to compare journals across disciplines (Moed, 2010).

### 2.3 Why Combining?

To retain the advantages of exploiting different indices and provide a unique, aggregate ranking, while overcoming the inherent limitations of using a single index, several meta-indices (*i.e.*, indices that combine a number of existing, base indices) have been defined. To date, a number of studies have focused on the comparison between rankings provided by the IF and the H-Index for journals in a certain field of study. Vanclay (2008) has reported data for IF and H-Index for forestry journals. In social work, Hodge and Lacasse (2010) have shown that the IF, the 5-year IF, the H-Index and the experts' opinion are correlated. Other studies have attempted to combine the IF and the H-Index with the aim of obtaining a meta-ranking. In particular, Bador and Lafouge (2010) have considered four groups of journals divided per quartile according to their categorical combined score in IF and H-Index. The underlying assumption in that work is that the number of citations per article is a random variable following a Paretian distribution with finite expectation. More recently, optimization-based consensus ranking has been exploited to construct suitable aggregates of individual journal rankings, considering journals from the Harzing List (Theu$\beta$l et al, 2014). As the authors discuss in their work, this method is however not very stable as the number of combined rankings (and list of journals) grows. Another combination method, named CombSUM, has been recently proposed to re-rank journals in computer science, by combining the IF and the H-Index (Tsai, 2014). The main disadvantage of this approach is that it does not allow one to assign different weights to the base indices, therefore reducing the flexibility of the combination scheme. Finally, Tüselmann et al (2015) have tried to shed more light on the problem of combining indices in the presence of *missing values*. They have proposed a modified Data Envelopment Analysis (DEA) model to estimate the missing values, and then exploited a statistical machine-learning approach (in particular, a *random forest*) to create a meta-index. However, a disadvantage is that the final meta-ranking is ambiguous, as it is characterized by several ties (this may be due to the choice of the random forest as the learning algorithm, as it naturally outputs discretized values).

In the next section, we propose a statistical approach that aims to overcome these limitations, by enabling us to combine an arbitrary number of indices, to assign them different weights, and to avoid making any specific underlying assumption on the data distribution.

## 3 Learning Aggregated Indices for Meta-Ranking

In this section, we discuss a methodology to aggregate existing indices, aiming to capture the different dimensions characterizing aspects of research *quality* in a consistent manner. Although this model can be also used to combine the experts' opinion, in this paper, we limit our analysis to the combination of quantitative indicators. As mentioned in Sect. 2, our goal is to propose an index aggregation scheme that overcomes the limitations emerged from the state of the art. To this end, we consider a simple linear combination of indices whose weights can be determined based on specific (and potentially different) criteria. Note also that some of the previously-proposed meta-indices can be expressed in terms of a linear combination of indices, as discussed in the following.

Let us assume we are given a set of journals $\mathcal{D} = \{\boldsymbol{x}_i\}_i^n$, where $\boldsymbol{x}_i = (x_i^1, \ldots, x_i^d) \in \mathbb{R}^d$ represent $d$ different index values for the $i^{\text{th}}$ journal; *e.g.*, $x_1^1$ and $x_1^2$ may respectively represent the H-index and the SJR for the first journal in the set $\mathcal{D}$. Our goal is then to learn an aggregated index as:

$$f(\boldsymbol{x}) = \sum_{k=1}^{d} w^k x^k + b \tag{2}$$

where $\boldsymbol{w} = (w^1, \ldots, w^d) \in \mathbb{R}^d$ is the $d$-dimensional vector of weights, each assigned to a different index, and $b$ is a bias, to allow $f$ to have a non-zero mean.[1] Different techniques can be exploited to learn $\boldsymbol{w}$ and $b$ in the above scheme. For instance, one is the DEA model proposed by Tüselmann et al (2015), which learns a set of weights $\boldsymbol{w}$, while using a null bias $b$ (see also Sect. 2.3). Furthermore, simple aggregation rules like CombSUM (Tsai, 2014) and Borda Count (specific for ranking) can be expressed in terms of this representation by assuming uniform weights (*i.e.*, $w^k = 1$, for $k = 1, \ldots, d$), and normalized index values. In particular, for the CombSUM method, indices may be normalized using min-max or Z normalization, respectively as:

$$x_i'^k = \frac{x_i^k - \min_{j=1,\ldots,n} x_j^k}{\max_{j=1,\ldots,n} x_j^k - \min_{j=1,\ldots,n} x_j^k}\,, \tag{3}$$

$$x_i'^k = \frac{x_i^k - \mu^k}{\sigma^k}\,, \tag{4}$$

where $x_i'^k$ is the normalized value for the $k^{\text{th}}$ index of the $i^{\text{th}}$ journal, and $\mu^k$ and $\sigma^k$ are the mean and standard deviation for the $k^{\text{th}}$ index values of the

[1] Note that, although the value of $b$ is irrelevant when ranking journals according to $f(\boldsymbol{x})$, it may be helpful during the process of learning the weights $\boldsymbol{w}$, as the values of the considered indices do not typically have zero mean.

journals in $\mathcal{D}$. For Borda Count, and similar ranking aggregation methods, we should consider as values of $x_i'^k$ the position of the $i^{\text{th}}$ journal in the ranked list of the $k^{\text{th}}$ index; in particular, if we are given $n = 100$ journals, and the $i^{\text{th}}$ journal is ranked $r = 5^{\text{th}}$ using the $k^{\text{th}}$ index, then $x_i'^k = n - r + 1 = 96$.

In general, to learn a linear combination function $f(\boldsymbol{x})$, *i.e.*, its parameters $\boldsymbol{w}$ and $b$, we are not restricted to the use of DEA or simple combination rules as the aforementioned ones. A set of different existing techniques proposed in the area of statistical data mining and machine learning can be exploited to this end (Bishop, 2007). For instance, one may project the data $\mathcal{D}$ onto a reduced vector space using Principal Component Analysis (PCA), and consider as the weights $\boldsymbol{w}$ the values of the first component (eigenvector). This will capture the direction of the vector space along which data is maximally spread (*i.e.*, exhibiting the highest variance). PCA is an example of an *unsupervised* learning technique, as it projects data onto a subspace without exploiting any knowledge of a desired *target* value. Conversely, *supervised* learning techniques assume that, for each sample in $\mathcal{D}$, we are also given a target value $y_i$, and learn $f$ by minimizing a functional of the form:

$$\min_{\boldsymbol{w},b} \frac{1}{n} \sum_{i=1}^{n} \ell\left(y_i, f(\boldsymbol{x}_i)\right) + \lambda \Omega(\boldsymbol{w}), \tag{5}$$

where $\ell\left(y_i, f(\boldsymbol{x}_i)\right)$ is a loss function that penalizes values of $f(\boldsymbol{x}_i)$ which are different from the target value $y_i$, $\Omega(\boldsymbol{w})$ is a regularization term that penalizes high values of $\boldsymbol{w}$ to provide a more stable solution, and $\lambda$ is a trade-off parameter. To be more concrete, let us give some examples. If we consider $\ell\left(y_i, f(\boldsymbol{x}_i)\right) = (y_i - f(\boldsymbol{x}_i))^2$, without regularization, we yield the classical minimum mean square error (MMSE) linear regression problem. If we consider an additional regularization term $\Omega(\boldsymbol{w}) = \|\boldsymbol{w}\|_2^2 = \sum_{k=1}^{d} \left(w^k\right)^2$ (*i.e.*, the $\ell_2$-norm of $\boldsymbol{w}$), and $\lambda > 0$, we yield *ridge* regression.

**Support Vector Regression.** Another very popular regression technique is Support Vector Regression (SVR) (Vapnik, 1995). It minimizes a functional as that given in Eq. (5), where $\ell\left(y_i, f(\boldsymbol{x}_i)\right) = \max\left(0, |y_i - f(\boldsymbol{x}_i)| - \varepsilon\right)$ is the so-called $\varepsilon$-insensitive loss, and $\Omega(\boldsymbol{w}) = \|\boldsymbol{w}\|_2^2$. This essentially assigns a linear penalty to points for which $y$ falls outside of a "tolerance" band $[f(\boldsymbol{x}) - \varepsilon, f(\boldsymbol{x}) + \varepsilon]$, as shown in Fig. 1. This technique can also be used to perform nonlinear regression tasks, by means of the so-called *kernel* trick, which allows one to write the function $f(\boldsymbol{x})$ as a linear combination of similarities (*i.e.*, kernel functions) computed between $\boldsymbol{x}$ and the so-called *support vectors* (*i.e.*, a subset of the training points in $\mathcal{D}$). This is why this technique is named support vector regression. As we focus on linear functions in this work, it is clearly out of the scope to provide further details on the use of nonlinear kernels here, for which we refer the reader to Vapnik (1995) and Bishop (2007).

**Defining the target values.** In our case, assuming known values of $y$ is equivalent to assuming that an *ideal* value of our aggregated index is already known for the journals in $\mathcal{D}$, which is clearly not the case (and it is indeed

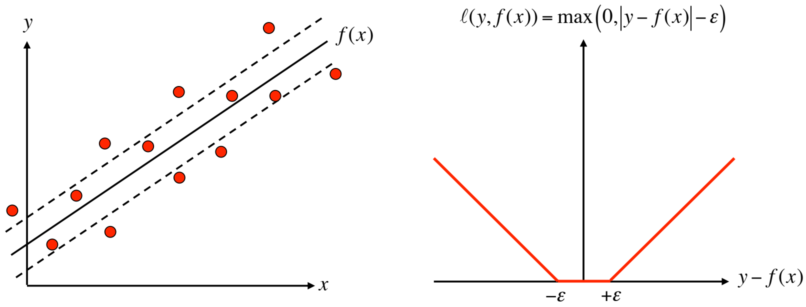


**Fig. 1** *Left:* SVR finds a linear function $f(\boldsymbol{x})$ that only penalizes $y$ values outside of the tolerance band $[f(\boldsymbol{x}) - \varepsilon, f(\boldsymbol{x}) + \varepsilon]$ (dashed black lines). *Right:* The $\varepsilon$-insensitive loss.

what we would like to achieve). However, we can somehow approximate the distribution of this value and make inference on that to estimate $\boldsymbol{w}$ and $b$ as discussed above, using the many supervised learning techniques that have already been proposed. To approximate $y$, we leverage on a similar idea to that exploited by Bador and Lafouge (2010). It amounts to producing a tied rank of journals by letting each single index vote whether a given journal should be in the first, second, third, or fourth quartile of the final distribution, and then summing up votes coming from the different available indices. Let us thus assume that the $k^{\text{th}}$ index expresses a vote $q^k \in \{0, \ldots, m-1\}$, where $m-1$ denotes the highest ranking (*e.g.*, the first quartile would correspond to $q = 3$, as, for quartiles, $m = 4$). Then the value $y_i$ for the $i^{\text{th}}$ journal is simply given by $y_i = \sum_{k=1}^{d} q^k$. The corresponding $y_i$ values will thus be in the set $\{0, \ldots, d(m-1)\}$, and several journals will have the same value of $y$, *i.e.*, the same rank in the final list. These ties can be broken to obtain an unambiguous final ranking by learning $\boldsymbol{w}$ and $b$ using one of the aforementioned supervised learning techniques; *e.g.*, ridge regression and SVR. The regularization parameter $\lambda$ in Eq. (5) is often estimated through a $k$-fold cross validation on the available data $\mathcal{D}$, to optimize performance while minimizing the risk of *overfitting*, *i.e.*, of learning functions that predict the training data with almost no error, but do not properly generalize on unseen data. This has to be especially accounted for in high-dimensional spaces, and when learning nonlinear functions. In our case, we exploit cross-validation to tune the parameter $\lambda$ by testing different values and retaining the one that minimizes the mean absolute error. We then retain the score $f(\boldsymbol{x})$ given by our method trained with the best value of $\lambda$ to the points in the validation fold.

## 4 Meta-Ranking of Business and Management Journals

In this section, we apply our analysis of journal meta-ranking using the two most important citation databases, *i.e.*, Thomson Reuters JCR and the Scopus SJR. We have selected two field of study: Business and Management. In building our datasets, we have decided to use the 5-year Impact Factor and the H-index to capture the stability, IF and SJR to capture the current trend. This mix of indices seems to be appropriate as SJR is closer to the size dimension than the IF, while H-Index attempts to capture both size and impact dimensions (Leydesdorff, 2009; Hodge and Lacasse, 2010).As pointed out in previous studies, the H-Index correlates highly with the Thomson Reuters 5-year Impact Factor and its scores are similar to the experts' opinion (Hodge and Lacasse, 2010). As already stated, SJR contains a large number of journals compared to JCR. Hence, to have consistent data, we have decided to use the Thomson dataset as "master". In particular, from that report, we have extracted IF and 5-year IF indices, whereas we have extracted the SJR and H-index from Scopus. In total, we have respectively $n = 173$ journals in Management and $n = 111$ in Business, corresponding to the whole journals indexed in Thomson Reuters for both subject areas. For both cases, we denote with $\mathcal{D} = \{\boldsymbol{x}_i\}_{i=1}^n$ the retrieved set of $n$ journals, where each journal $\boldsymbol{x}_i$ is a four-dimensional vector characterized by the four index values IF, 5-year IF, SJR, and H-index.

**Setup.** For both journals in Management and Business, we have aggregated the four baseline indices into the following meta-indices: SVR, CombSUM, Borda Count and PCA. To define the ground-truth labels $y$ required to train the SVR, we have used quartile-based voting for each baseline index, *i.e.*, we have set $m = 4$ (see Sect 3). The regularization parameter $\lambda \in \frac{1}{n}\{10^{-3}, 10^{-2}, \dots, 10^3\}$ of the SVR has been selected using a 5-fold cross-validation procedure aiming to maximize the mean absolute error on the validation fold.

**Results.** In Figs. 2 and 3, we show Kendall's $\tau$ correlations between each meta-index and the baseline indices. Note that this kind of correlation is more suitable for evaluating differences in rankings, as it depends more on the relative order of entries rather than the associated numeric index values. The considered indices are clearly correlated with each other, with correlation values higher than 0.7 for almost all index pairs, except for those involving the H-Index. Despite the correlation between baseline and meta-indices may seem rather high, the corresponding rankings may exhibit significant variations, as one may appreciate from Tables 1 and 3, where we report the top 20 journals according to the SVR index, and how they are ranked by the other indices. Besides capturing the multidimensionality of the quality concept according to slightly different facets, which will definitely need further investigation in the future (also considering the experts' opinions), this behavior raises the question of whether and to what extent meta-indices may be affected by noise in the base index values, *i.e.*, if they are able to provide a *stable* meta-ranking. This issue is investigated in the experiments of Sect. 4.1.

**Results with nine combined indices.** To provide additional insights on the selection of the indices to combine, we have considered five additional citation indices, namely, the Eigenfactor Score ('Eig'), the Article Influence Score (AIS), the Immediacy Index ('Imm'), IPP and SNIP. Exploiting the aforementioned techniques, we report in Tables 2 and 4 the corresponding rankings for

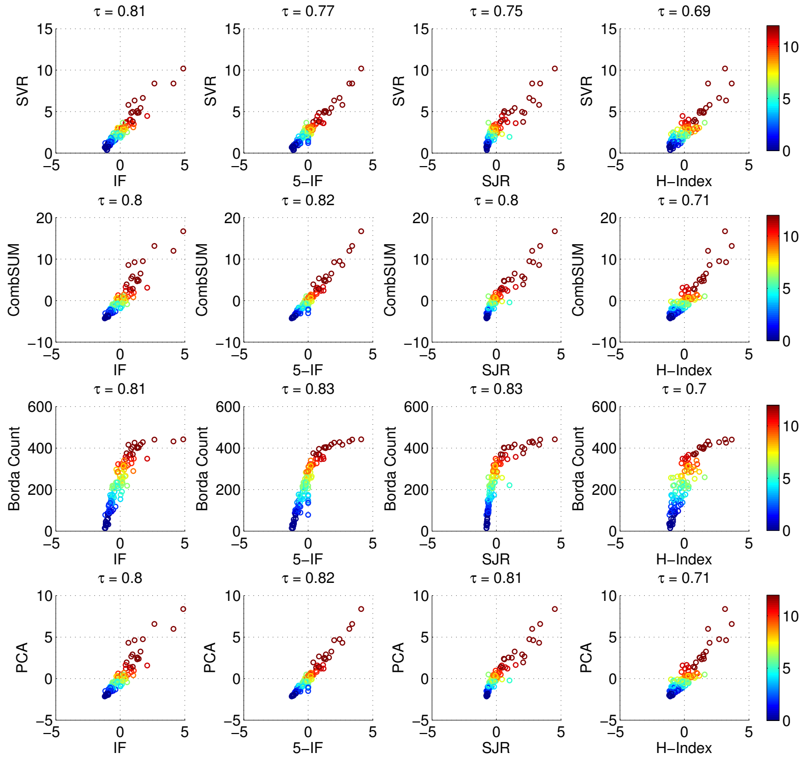


**Fig. 2** Kendall's $\tau$ correlations between each meta-index (rows) and each base index (columns) for journals in the Business area. Each point in the scatter plots represents a distinct journal, and its color denotes the corresponding ground-truth value $y$.

Business and Management journals. Although adding these indices does not bring radical changes in the overall ranking, we will see how combining more base indices can improve the stability of meta-ranking against some random (noisy) perturbations of the values of the base indices (Sect. 4.1).

**Visual Index Comparison.** We finally consider a different projection to visualize each index in a compact, two-dimensional vector space, and evaluate again how similar they are to each other, similarly to the procedure adopted by Leydesdorff (2009). In particular, we apply again PCA, but this time considering each index as a point, and the values it assigns to each journal as its dimensions. The first two principal components of this projection are shown in Fig. 4, where it can be appreciated how almost all meta-indices (except for Borda Count) are close to each other and well-summarize the characteristics of the four combined base indices. Kendall's $\tau$ correlations (computed in the non-reduced space, using all $n$ journals as dimensions) are also reported in Table 4 for the sake of completeness. As expected, also this analysis highlights that the considered meta-indices remain similar even if the number of combined base indices increases, properly summarizing their characteristics.

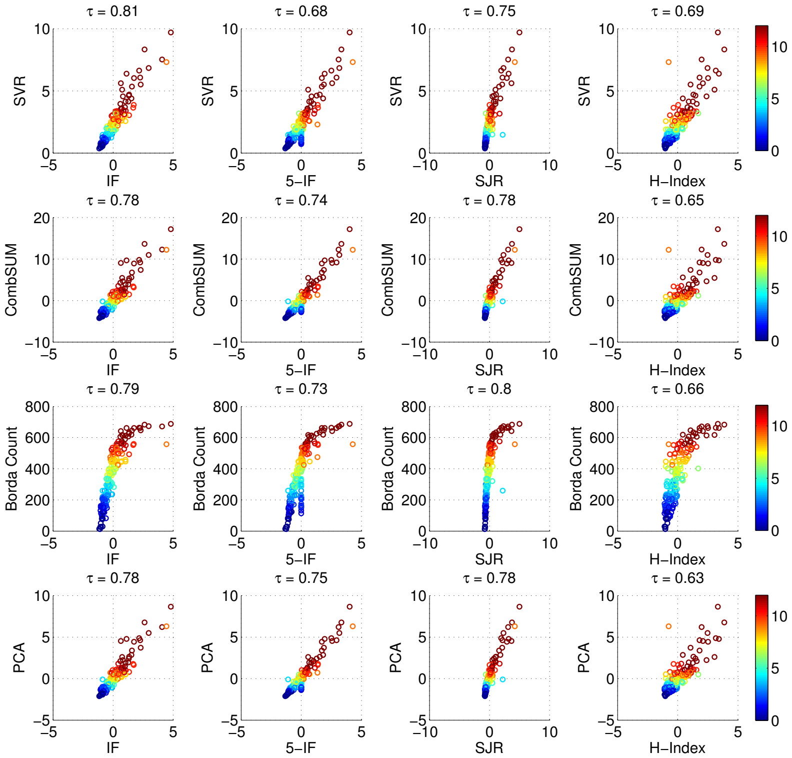


**Fig. 3** Kendall's $\tau$ correlations between each meta-index (rows) and each base index (columns) for journals in the Management area. See the caption of Fig. 2 for further details.

## 4.1 Meta-Ranking Stability against Perturbation of the Base Index Values

In this section we report a simple analysis to evaluate the stability of the proposed meta-indices under perturbation of the values of the base indices. This may be of interest as one may deal with noisy data, or some index values may not be completely reliable for some journals (*e.g.*, in the case of missing data, or missing citations, *etc.*). Accordingly, one may be interested in knowing whether meta-indices enjoy some robustness property against noise in the input data, or if they are significantly affected. In the latter case, in fact, meta-ranking may be of little practical impact.

For the sake of the analysis presented in this section, we have simulated a random Gaussian noise (with zero mean and standard deviation $\sigma$) over the base index values, after having normalized the latter to have zero mean and unary standard deviation. Results for an increasing level of noise (*i.e.*, increasing noise standard deviation $\sigma$) are reported in Fig. 5. To evaluate stability of meta-rankings, we have measured the Kendall's $\tau$ correlation between the ranking obtained in the absence of noise ($\sigma = 0$) and that obtained in the presence of noise in the input index values ($\sigma > 0$) by the same meta-index.

| Journal / Index | IF | 5-IF | SJR | H | SVR | CS | BC | PCA |
|---|---|---|---|---|---|---|---|---|
| Academy of Manag. Review | 1 | 1 | 1 | 3 | 1 | 1 | 1 | 1 |
| Academy of Manag. Journal | 3 | 2 | 2 | 1 | 2 | 2 | 2 | 2 |
| Journal of Manag. | 2 | 3 | 5 | 6 | 3 | 3 | 3 | 3 |
| Journal of Marketing | 5 | 5 | 6 | 4 | 4 | 4 | 4 | 4 |
| Strategic Manag. Journal | 11 | 6 | 4 | 2 | 5 | 5 | 5 | 5 |
| Adm. Science Quarterly | 21 | 4 | 3 | 5 | 6 | 6 | 7 | 6 |
| Journal of Int'l Business Studies | 6 | 7 | 11 | 7 | 7 | 7 | 6 | 7 |
| Journal of Business Venturing | 9 | 11 | 12 | 13 | 8 | 11 | 10 | 11 |
| Journal of Org'l Behavior | 10 | 10 | 17 | 12 | 9 | 13 | 12 | 13 |
| Journal of Consumer Research | 15 | 9 | 9 | 10 | 10 | 8 | 8 | 8 |
| J. Academy Market. Science | 7 | 12 | 18 | 9 | 11 | 12 | 11 | 12 |
| Journal of Marketing Research | 18 | 17 | 7 | 11 | 12 | 9 | 13 | 9 |
| Journal of Manag. Studies | 8 | 8 | 14 | 14 | 13 | 10 | 9 | 10 |
| Family Business Review | 4 | 19 | 21 | 54 | 14 | 16 | 19 | 16 |
| Entrepr. Theory and Practice | 19 | 16 | 16 | 29 | 15 | 18 | 17 | 18 |
| J. Env. Economics and Manag. | 20 | 21 | 19 | 21 | 16 | 19 | 16 | 19 |
| Academy of Manag. Persp. | 14 | 18 | 23 | 22 | 17 | 17 | 15 | 17 |
| Marketing Science | 24 | 24 | 8 | 18 | 18 | 14 | 14 | 14 |
| Journal of Service Research | 26 | 15 | 15 | 47 | 19 | 21 | 22 | 21 |
| Harvard Business Review | 37 | 56 | 88 | 8 | 20 | 26 | 45 | 26 |

**Table 1** Differences in ranking and meta-ranking for journals in the Business area. Journals are sorted here according to the SVR ranking. The corresponding rank according to each other index is reported in the corresponding column. IF and 5-IF stand for Impact Factor and 5-year IF, H for H-index, CS for CombSUM, and BC for Borda Count.

Accordingly, $\tau = 1$ when $\sigma = 0$ for all meta-indices, while it start decreasing as $\sigma$ increases. Clearly, the more stable rankings are those for which the correlation value $\tau$ decreases more gracefully with respect to $\sigma$. Interestingly, the reported results show that SVR is quite unstable in the presence of even very small noise in the input index values, whereas the other methods are more stable. This may be due to the fact that the introduced noise also affects the estimated ground-truth values $y$, and the SVR, being a supervised technique, strongly depends on them. It would thus be of interest in the future to understand how to overcome this limitation, by reducing the dependency of the SVR on the ground-truth values. Furthermore, another interesting observation is that combining more indices seems to improve stability. An intuitive explanation may be that *averaging* more indices may reduce the variance of the estimation error, similarly to the reduction of the variance of the estimation error of the sample mean with the increase of the number of averaged items.

To summarize, the rankings obtained with four and nine base indices seems to be quite similar, *i.e.*, they properly capture the multidimensionality of the quantitative aspect of journal quality, without biasing ranking towards any of the base dimensions. However, stability is improved when more base indices are combined, especially if proper meta-ranking methods are considered. Accordingly, our overall empirical analysis suggests that combining the given nine indices using either CombSUM or PCA may be more appropriate.

| Journal / Index | IF | 5-IF | SJR | H | Imm | Eig | AIS | SNIP | IPP | SVR | CS | BC | PCA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Academy of Manag. Review | 1 | 1 | 1 | 3 | 2 | 7 | 2 | 1 | 1 | 1 | 1 | 1 | 1 |
| J. of Manag. | 2 | 3 | 5 | 6 | 15 | 3 | 4 | 2 | 2 | 2 | 3 | 3 | 3 |
| Academy of Manag. J. | 3 | 2 | 2 | 1 | 21 | 1 | 3 | 5 | 3 | 3 | 2 | 2 | 2 |
| J. of Marketing | 5 | 5 | 6 | 4 | 36 | 6 | 5 | 3 | 4 | 4 | 4 | 4 | 4 |
| Strategic Manag. J. | 11 | 6 | 4 | 2 | 33 | 2 | 6 | 9 | 9 | 5 | 5 | 5 | 5 |
| Adm. Science Quarterly | 21 | 4 | 3 | 5 | 28 | 16 | 1 | 10 | 5 | 6 | 6 | 7 | 6 |
| J. of Int'l Business Studies | 6 | 7 | 11 | 7 | 7 | 11 | 15 | 13 | 10 | 7 | 7 | 6 | 8 |
| J. of Manag. Studies | 8 | 8 | 14 | 14 | 9 | 10 | 9 | 12 | 11 | 8 | 9 | 8 | 10 |
| J. of Business Venturing | 9 | 11 | 12 | 13 | 16 | 17 | 14 | 6 | 7 | 9 | 11 | 9 | 11 |
| J. of Consumer Research | 15 | 9 | 9 | 10 | 26 | 5 | 8 | 11 | 13 | 10 | 8 | 10 | 7 |
| J. of Organizational Behavior | 10 | 10 | 17 | 12 | 20 | 13 | 12 | 15 | 14 | 11 | 12 | 11 | 12 |
| Int'l J. of Manag. Reviews | 17 | 13 | 24 | 46 | 23 | 31 | 17 | 4 | 6 | 12 | 15 | 17 | 15 |
| J. Academy Market. Science | 7 | 12 | 18 | 9 | 55 | 20 | 19 | 14 | 12 | 13 | 14 | 15 | 14 |
| J. of Marketing Research | 18 | 17 | 7 | 11 | 32 | 4 | 7 | 17 | 16 | 14 | 10 | 12 | 9 |
| Academy of Manag. Persp. | 14 | 18 | 23 | 22 | 5 | 25 | 16 | 20 | 20 | 15 | 16 | 13 | 17 |
| Family Business Review | 4 | 19 | 21 | 54 | 8 | 48 | 24 | 27 | 19 | 16 | 18 | 18 | 19 |
| J. of Service Research | 26 | 15 | 15 | 47 | 70 | 34 | 25 | 8 | 8 | 17 | 19 | 20 | 18 |
| J. Env. Economics and Manag. | 20 | 21 | 19 | 21 | 31 | 15 | 13 | 16 | 23 | 18 | 17 | 16 | 16 |
| J. of Advertising Research | 47 | 69 | 71 | 42 | 1 | 52 | 60 | 89 | 88 | 19 | 22 | 58 | 42 |
| Marketing Science | 24 | 24 | 8 | 18 | 30 | 9 | 10 | 19 | 22 | 20 | 13 | 14 | 13 |

**Table 2** Differences in ranking and meta-ranking for journals in the Business area, when combining nine base indices. 'Imm', 'Eig' and 'AIS' denote the Immediacy Index, the Eigenfactor Score and the Article Influence Score. See caption of Tab. 1 for further details.

| Journal / Index | IF | 5-IF | SJR | H | SVR | CS | BC | PCA |
|---|---|---|---|---|---|---|---|---|
| Academy of Manag. Review | 1 | 2 | 1 | 3 | 1 | 1 | 1 | 1 |
| Academy of Manag. J. | 5 | 3 | 3 | 1 | 2 | 2 | 2 | 2 |
| J. of Manag. | 3 | 5 | 6 | 10 | 3 | 3 | 4 | 4 |
| Academy of Manag. Annals | 2 | 1 | 2 | 131* | 4 | 4 | 29 | 3 |
| MIS Quarterly: Manag. Inf. Sys. | 4 | 4 | 8 | 7 | 5 | 5 | 3 | 5 |
| J. of Applied Psychology | 8 | 8 | 12 | 4 | 6 | 6 | 5 | 6 |
| Strategic Manag. J. | 18 | 9 | 5 | 2 | 7 | 7 | 7 | 7 |
| J. of Operations Manag. | 7 | 6 | 10 | 12 | 8 | 8 | 6 | 8 |
| Organization Science | 9 | 13 | 7 | 6 | 9 | 10 | 8 | 10 |
| Adm. Science Quarterly | 34 | 7 | 4 | 9 | 10 | 9 | 11 | 9 |
| Personnel Psychology | 6 | 10 | 11 | 24 | 11 | 11 | 9 | 11 |
| J. of Int'l Business Studies | 12 | 12 | 15 | 11 | 12 | 12 | 10 | 12 |
| Manag. Science | 29 | 28 | 17 | 5 | 13 | 14 | 15 | 15 |
| J. of Manag. Studies | 14 | 14 | 16 | 18 | 14 | 15 | 12 | 14 |
| J. of Organizational Behavior | 16 | 18 | 21 | 16 | 15 | 16 | 13 | 16 |
| Research Policy | 27 | 22 | 26 | 8 | 16 | 17 | 16 | 17 |
| Org'l Research Methods | 13 | 11 | 13 | 42 | 17 | 13 | 14 | 13 |
| Org. Behavior Human Dec. Proc. | 20 | 23 | 22 | 19 | 18 | 20 | 17 | 20 |
| Omega | 17 | 26 | 19 | 28 | 19 | 19 | 18 | 19 |
| Information Systems Research | 37 | 21 | 18 | 14 | 20 | 18 | 19 | 18 |

**Table 3** Differences in ranking and meta-ranking for journals in the Management area, when combining four base indices. See caption of Tab. 1 for further details. *This journal has low rank for the H-index, as its data is available only from 2011.

## 5 Discussion and Conclusions

Producing a reliable, widely-approved journal ranking is a non-trivial task, mainly due to the inherent difficulty of selecting a proper set of base indices and combination technique (*i.e.*, meta-index) among the existing ones. In this paper, we have highlighted the need of combining various indices, by leveraging

| Journal / Index | IF | 5-IF | SJR | H | Imm | Eig | AIS | SNIP | IPP | SVR | CS | BC | PCA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Academy of Manag. Review | 1 | 2 | 1 | 3 | 1 | 9 | 3 | 1 | 1 | 1 | 1 | 1 | 1 |
| J. of Manag. | 3 | 5 | 6 | 10 | 18 | 6 | 5 | 4 | 3 | 2 | 4 | 3 | 4 |
| Academy of Manag. J. | 5 | 3 | 3 | 1 | 27 | 3 | 4 | 6 | 6 | 3 | 2 | 2 | 2 |
| MIS Quarterly: Manag. Inf. Sys. | 4 | 4 | 8 | 7 | 6 | 13 | 12 | 2 | 4 | 4 | 5 | 4 | 5 |
| Academy of Manag. Annals | 2 | 1 | 2 | 131* | 21 | 24 | 1 | 3 | 2 | 5 | 3 | 14 | 3 |
| J. of Applied Psychology | 8 | 8 | 12 | 4 | 38 | 2 | 7 | 9 | 9 | 6 | 6 | 5 | 6 |
| J. of Operations Manag. | 7 | 6 | 10 | 12 | 29 | 26 | 17 | 7 | 5 | 7 | 10 | 7 | 10 |
| Adm. Science Quarterly | 34 | 7 | 4 | 9 | 37 | 18 | 2 | 10 | 7 | 8 | 8 | 10 | 7 |
| Strategic Manag. J. | 18 | 9 | 5 | 2 | 45 | 5 | 10 | 8 | 11 | 9 | 7 | 6 | 8 |
| Personnel Psychology | 6 | 10 | 11 | 24 | 39 | 21 | 9 | 11 | 10 | 10 | 11 | 12 | 11 |
| J. of Int'l Business Studies | 12 | 12 | 15 | 11 | 11 | 11 | 18 | 16 | 12 | 11 | 12 | 8 | 12 |
| Organization Science | 9 | 13 | 7 | 6 | 56 | 4 | 6 | 19 | 14 | 12 | 9 | 11 | 9 |
| J. of Manag. Studies | 14 | 14 | 16 | 18 | 12 | 10 | 15 | 13 | 13 | 13 | 13 | 9 | 13 |
| Omega | 17 | 26 | 19 | 28 | 5 | 23 | 42 | 12 | 19 | 14 | 18 | 15 | 18 |
| Int'l J. of Manag. Reviews | 26 | 19 | 33 | 55 | 30 | 48 | 24 | 5 | 8 | 15 | 19 | 21 | 19 |
| J. of Organizational Behavior | 16 | 18 | 21 | 16 | 24 | 15 | 16 | 18 | 15 | 16 | 15 | 13 | 16 |
| Research Policy | 27 | 22 | 26 | 8 | 51 | 8 | 26 | 15 | 18 | 17 | 17 | 16 | 17 |
| Organizational Research Methods | 13 | 11 | 13 | 42 | 61 | 17 | 11 | 24 | 16 | 18 | 16 | 17 | 15 |
| Academy of Manag. Persp. | 23 | 24 | 30 | 30 | 9 | 37 | 21 | 26 | 24 | 19 | 21 | 19 | 22 |
| Human Res. Manag. Review | 41 | 36 | 45 | 52 | 4 | 52 | 34 | 14 | 23 | 20 | 23 | 24 | 25 |

**Table 4** Differences in ranking and meta-ranking for journals in the Management area, when combining nine base indices. See caption of Tab. 1 for further details. *This journal has low rank for the H-Index, as its data is available only from 2011.

| | SVR | CS | BC | PCA |
|---|---|---|---|---|
| IF | 0.81 | 0.80 | 0.81 | 0.80 |
| 5-IF | 0.77 | 0.82 | 0.83 | 0.82 |
| SJR | 0.75 | 0.80 | 0.83 | 0.81 |
| H-Index | 0.69 | 0.71 | 0.70 | 0.71 |
| Avg.: | 0.76 | 0.78 | 0.79 | 0.78 |
| Std.: | 0.05 | 0.05 | 0.06 | 0.05 |

| | SVR | CS | BC | PCA |
|---|---|---|---|---|
| IF | 0.81 | 0.78 | 0.79 | 0.78 |
| 5-IF | 0.68 | 0.74 | 0.73 | 0.75 |
| SJR | 0.75 | 0.78 | 0.80 | 0.78 |
| H-Index | 0.69 | 0.65 | 0.66 | 0.63 |
| Avg.: | 0.73 | 0.74 | 0.74 | 0.74 |
| Std.: | 0.06 | 0.06 | 0.07 | 0.07 |

| | SVR | CS | BC | PCA |
|---|---|---|---|---|
| IF | 0.79 | 0.77 | 0.77 | 0.77 |
| 5-IF | 0.76 | 0.76 | 0.79 | 0.78 |
| SJR | 0.76 | 0.78 | 0.82 | 0.80 |
| H-Index | 0.69 | 0.71 | 0.70 | 0.71 |
| Imm | 0.56 | 0.58 | 0.56 | 0.54 |
| Eig | 0.72 | 0.76 | 0.76 | 0.77 |
| AIS | 0.69 | 0.71 | 0.74 | 0.73 |
| SNIP | 0.74 | 0.73 | 0.73 | 0.74 |
| IPP | 0.81 | 0.79 | 0.80 | 0.80 |
| Avg.: | 0.73 | 0.73 | 0.74 | 0.74 |
| Std.: | 0.07 | 0.06 | 0.08 | 0.08 |

| | SVR | CS | BC | PCA |
|---|---|---|---|---|
| IF | 0.79 | 0.77 | 0.78 | 0.78 |
| 5-IF | 0.68 | 0.72 | 0.73 | 0.74 |
| SJR | 0.73 | 0.77 | 0.78 | 0.77 |
| H-Index | 0.64 | 0.61 | 0.63 | 0.61 |
| Imm | 0.55 | 0.54 | 0.54 | 0.52 |
| Eig | 0.69 | 0.71 | 0.74 | 0.71 |
| AIS | 0.58 | 0.65 | 0.65 | 0.65 |
| SNIP | 0.76 | 0.73 | 0.73 | 0.74 |
| IPP | 0.81 | 0.76 | 0.77 | 0.77 |
| Avg.: | 0.69 | 0.70 | 0.70 | 0.70 |
| Std.: | 0.09 | 0.08 | 0.08 | 0.09 |

**Table 5** Kendall's $\tau$ correlation between meta-indices and base indices for Business (left) and Management (right) journals, when combining four (top row) and nine (bottom row) base indices. The average correlation (along with the corresponding standard deviation) between each meta-index and the base indices is also reported.

on the main limitations emerged from previous work. We have proposed and formalized an approach able to reliably rank journals, and to capture the different dimensions characterizing the aspects of research *quality* in a consistent manner. Firstly, we have determined if a given journal should be in the first, second, third, or fourth quartile of the final distribution, according to each index. Secondly, using different techniques, we have aggregated the votes coming from the different base indices. Finally, we have sorted journals for both

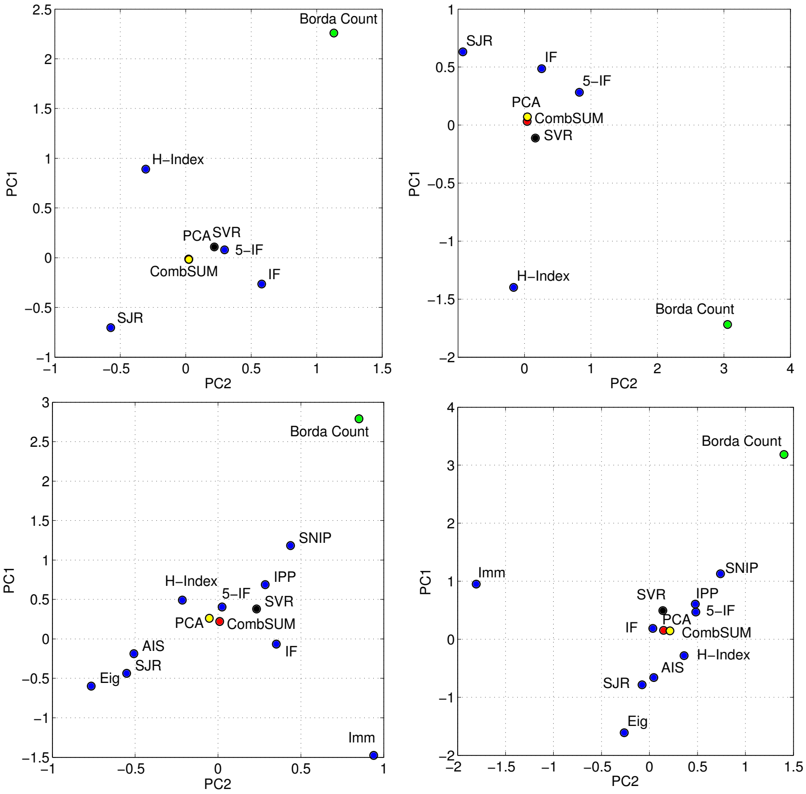


**Fig. 4** PCA-based projection on a two-dimensional space for Business (left column) and Management (right column) journals, when combining four (top row) and nine (bottom row) base indices. Each point is an index in the space of the first two principal components. Note that PCA and CombSUM meta-indices are overlapped in the top-left plot.

Business and Management area, according to the considered indices and meta-indices. We have also evaluated the performance of each meta-index, finding a high correlation between indices and meta-indices. Moreover, we have evaluated their distance in a two-dimensional vector space to visualize how similar they are to each other. In order to complete our analysis, we have provided a stability analysis against random noise in the databases, finding that the combination of nine selected indices improves stability in comparison to a smaller number of base indices, without affecting the corresponding meta-rankings. Our analysis has shown that both supervised and unsupervised learning techniques are all qualified tools to produce aggregate indices for journal ranking, despite supervised techniques may be more sensitive to noise in the input data.

Although we have chosen the combined indices according to a well-motivated principle to balance the contribution of stability and of the current trend, it is still an open issue to quantitatively evaluate how and to what extent the proposed meta-indices can be retained properly representative of the afore-

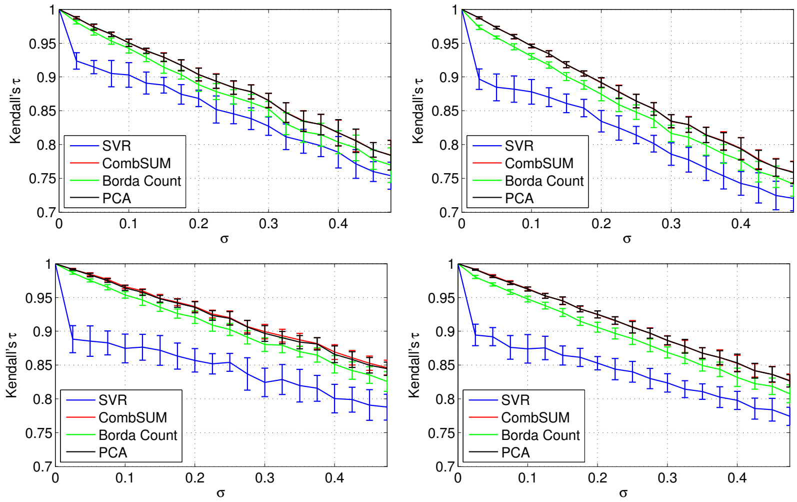


**Fig. 5** Stability analysis of meta-indices for Business (left column) and Management (right column) journals, when combining four (top row) and nine (bottom row) base indices.

mentioned aspects. To this end, we envision the possibility of combining and comparing the proposed technique with qualitative approaches (*e.g.*, based on the analysis of experts' opinions). This can be definitely considered a promising research direction. As previous work has been mainly focused on defining novel combination methods to aggregate a set of given base indices, there is need of shedding more light on how to select a proper set of indices to be combined, also taking into account the given combination method. This is another relevant research direction that may be worth investigating in the future. We finally believe that our framework can provide useful results for many purposes, *e.g.*, for researchers, as a reference to choose their publication outlets, and for faculties, departments and editors to evaluate and compare the quality of their own journals.

## References

Bador P, Lafouge T (2010) Comparative analysis between impact factor and h-index for pharmacology and psychiatry journals. Scientometrics 84(1):65–79

Bergstrom C (2007) Measuring the value and prestige of scholarly journals. College & Research Libraries News 68(5):314–316

Bishop CM (2007) Pattern Recognition and Machine Learning (Information Science and Statistics), 1st edn. Springer

Bornmann L, Marx W, Gasparyan AY, Kitas GD (2012) Diversity, value and limitations of the journal impact factor and alternative metrics. Rheumatology international 32(7):1861–1867

Braun T, Glänzel W, Schubert A (2006) A hirsch-type index for journals. Scientometrics 69(1):169–173

DuBois FL, Reeb D (2000) Ranking the international business journals. Journal of International Business Studies 31(4):689–704

Eigenfactor (2015) - scimago journal & country rank. University of Washington Retrieved in July 2015, URL `http://www.eigenfactor.org/about.php`

Falagas ME, Kouranos VD, Arencibia-Jorge R, Karageorgopoulos DE (2008) Comparison of scimago journal rank indicator with journal impact factor. The FASEB journal 22(8):2623–2628

Garfield E (2006) Citation indexes for science. a new dimension in documentation through association of ideas. International Journal of Epidemiology 35(5):1123–1127

Garfield E, Sher IH (1963) New factors in the evaluation of scientific literature through citation indexing. American documentation 14(3):195–201

González-Pereira B, Guerrero-Bote VP, Moya-Anegón F (2010) A new approach to the metric of journals' scientific prestige: The sjr indicator. Journal of informetrics 4(3):379–391

Hirsch JE (2005) An index to quantify an individual's scientific research output. Proceedings of the National academy of Sciences of the United States of America 102(46):16,569–16,572

Hodge DR, Lacasse JR (2010) Evaluating journal quality: Is the h-index a better measure than impact factors? Research on Social Work Practice

Lancho-Barrantes BS, Guerrero-Bote VP, Moya-Anegón F (2010) What lies behind the averages and significance of citation indicators in different disciplines? Journal of Information Science 36(3):371–382

Leydesdorff L (2009) How are new citation-based journal indicators adding to the bibliometric toolbox? Journal of the American Society for Information Science and Technology 60(7):1327–1336

Leydesdorff L, de Moya-Anegón F, Guerrero-Bote VP (2010) Journal maps on the basis of scopus data: A comparison with the journal citation reports of the isi. Journal of the American Society for information Science and Technology 61(2):352–369

Mingers J, Leydesdorff L (2015) A review of theory and practice in scientometrics. European Journal of Operational Research

Moed HF (2010) The source-normalized impact per paper (snip) is a valid and sophisticated indicator of journal citation impact. arXiv preprint arXiv:10054906

Mylonopoulos NA, Theoharakis V (2001) On site: global perceptions of is journals. Communications of the ACM 44(9):29–33

Peffers K, Ya T (2003) Identifying and evaluating the universe of outlets for information systems research: Ranking the journals. Journal of Information Technology Theory and Application (JITTA) 5(1):6

Scopus (2015) Journal metrics. SCimago Retrieved in July 2015, URL `http://www.journalmetrics.com`

Seglen PO (1997) Why the impact factor of journals should not be used for evaluating research. Bmj 314(7079):497

Sellers SL, Mathiesen SG, Perry R, Smith T (2004) Evaluation of social work journal quality: Citation versus reputation approaches. Journal of Social Work Education 40(1):143–160

SJR (2007) - scimago journal & country rank. SCimago Retrieved in February 2015, URL http://www.scimagojr.com

Theu$\beta$l S, Reutterer T, Hornik K (2014) How to derive consensus among various marketing journal rankings? Journal of Business Research 67(5):998–1006

Tsai CF (2014) Citation impact analysis of top ranked computer science journals and their rankings. Journal of Informetrics 8(2):318 – 328

Tüselmann H, Sinkovics RR, Pishchulov G (2015) Towards a consolidation of worldwide journal rankings–a classification using random forests and aggregate rating via data envelopment analysis. Omega 51:11–23

Vanclay JK (2008) Ranking forestry journals using the h-index. Journal of informetrics 2(4):326–334

Vapnik VN (1995) The nature of statistical learning theory. Springer-Verlag New York, Inc., New York, NY, USA

Vieira ES, Gomes JA (2009) A comparison of scopus and web of science for a typical university. Scientometrics 81(2):587–600

Vokurka RJ (1996) The relative importance of journals used in operations management research a citation analysis. Journal of Operations Management 14(4):345–355