Bibliometrics effects of a new paper level classification

Marcos Peña-Rocha[1], Rocío Gómez-Crisóstomo[1],
Vicente P. Guerrero-Bote[1]* and Félix de Moya-Anegón[2]

[1]Departamento de Información y Comunicación, Universidad de Extremadura, Badajoz, Spain,
[2]SCImago Research Group, Granada, Spain

This study presents a comparative analysis between two scientific document classification systems. The first system employs the Scopus journal-based assignment method, adapted to a fractional model, while the second system uses an item-by-item system based on reclassified references according to the origin of the citers. The study's results are divided into three different sections: the first involves comparisons at the Scopus area level, the second examines comparisons at the category level, and the third tests various bibliometric indicators to identify the variations between the two systems. Highlighting the characteristics of the paper level system, it offers a reduction in the number of categories to which each document is assigned, achieving higher values of single-category assignment compared to the All Science Journal Classification (ASJC). When reclassifying areas and categories, the paper level system tends to accentuate differences at the extreme values, increasing the size of the largest categories and reducing that of the smallest ones. Moreover, the paper-by-paper system provides more homogeneous distributions in normalised impacts and adjusts values related to excellence more uniformly.

1 Introduction

A widely accepted classification by the scientific community is essential both to determine and quantify the progress made in each discipline and to normalise the different citation habits that exist among them.

Currently, the classifications used by bibliographic databases such as Web of Science (WoS) or Scopus are the most widely used for these purposes. These classifications operate at journal-level, where each paper is assigned to the category to which the journal in which it is published is ascribed.

Classifying documents based solely on the category assigned to a journal has some drawbacks that generate inaccuracies in bibliometric studies. One such drawback is associated with the multiple category assignments that journals have. As Álvarez-Llorente et al. (2023) state, "*not all the work a journal publishes are from all the categories to which it is assigned, indeed, quite the contrary is the case*", causing inaccuracies since works are assigned to all the categories of the journal even though most belong to only one of them. This imprecision can even extend to both the classification of works and the normalisation of the scientific indicators previously mentioned.

Related to the previous issue, another inconvenience arises due to multidisciplinary categories. Journals classified under these categories publish works from various disciplines, but not all of their works are multidisciplinary. This situation creates a particular need to determine the specific disciplines to which each work belongs.

The initial attempts to address this aforementioned issue (Glänzel et al., 1999a,b) focused on the reference lists provided by the documents to classify them, indicating that references are the best way to define the content of a work, despite the different motivations for citing other works (Bornmann and Daniel, 2008). In 2021, Glänzel et al. proposed an improvement over the previous method by using a parameterised model that works with multiple generations of references (Glänzel et al., 2021).

In 2003, Glänzel and Schubert attempted to develop a classification system adapted for bibliometric use. In their study (Glänzel and Schubert, 2003), they proposed a hierarchical two-level classification system, developed through a three-step iterative process.

Years later, the implementation of clustering algorithms led to new methods for creating scientific classification schemes. By the late 2000s, several studies emerged based on hybrid clustering algorithms.

In 2008, Janssens et al. sought to classify the field of information science into different clusters (Janssens et al., 2008). In a subsequent study, they used an analysis that combined citations with the full text of the article to classify WoS journals for the period 2002–2006 (Janssens et al., 2009). They concluded that using the combination of both components yielded more favourable results than treating them separately.

Boyack and Klavans (2010) tested the accuracy of several clustering algorithms on a large set of biomedical documents. The selected algorithms used co-citation analysis, bibliographic coupling, direct citations, and the combination of citations with the document's text to generate clusters. The last approach, which used a hybrid method based on bibliographic coupling, stood out by offering better results than the others.

The following year, another comparative study was conducted (Boyack et al., 2011) with a sample of over two million MEDLINE documents. This study compared the accuracy of nine clustering algorithms based on text similarity between documents with the information extracted from their bibliographic records. Using document titles and abstracts on one hand and MeSH thesaurus terms on the other, they developed a summary table that grouped the results generated by the different algorithms. They state that the indicator known as PMRA (PubMed Related Articles), described in 2007 (Lin and Wilbur, 2007) generates the most coherent clusters with a low computational cost. This indicator is used by the PubMed database to recommend articles related to the document selected by the users. Its functionality is based on the Poisson probabilistic distribution, using both MeSH thesaurus terms and the words from the abstracts and titles, with title words given twice the weight of those in the abstracts.

In 2012, researchers Waltman and van Eck (2012) conducted a study in which they presented a methodology for classifying ~10 million documents without requiring high computational cost, using direct citations as a method to determine the relationship between works. They concluded that using other types of bibliographic relationships, such as bibliographic coupling, could offer significant improvements. Additionally, they identified several issues to be addressed in future research, such as improving the labelling of areas, allowing publications to be assigned to more than one area, and incorporating new articles into an existing classification.

In another attempt to improve cluster precision, Boyack et al. (2013) sought to enhance the accuracy of co-citation clustering models by using a proximity analysis of references from a total of 270,521 full-text documents from 2007. After comparing the traditional co-citation model with the hybrid method, they concluded that the hybrid model achieves greater precision and reduces the size of the generated clusters. This finding is consistent with their previous studies, where hybrid algorithms produced more accurate results.

Although clustering algorithms provide solutions for classification needs, using the classifications they generate has significant drawbacks compared to established classifications. These algorithms create structures based on the papers available at the time of execution. Therefore, when incorporating new papers, it becomes necessary to run the algorithm again, which can result in a completely different structure and very different groupings. Additionally, there is the issue of labelling clusters, which is far from trivial.

For these reasons, none of the previous clustering proposals have been widely accepted by the scientific community. Consequently, in recent years, there have been several attempts to classify documents using schemes that are accepted by the scientific community, such as those outlined below.

Milojević (2020) aimed to reclassify WoS using the classification scheme employed by the database. Milojević's reclassification work involved a total of 45 million documents, consisting of conference proceedings and scientific articles, each with at least one reference. The proposed method seeks to infer the category of an article through its references, with the goal of assigning documents to a single category. The first iteration of the system starts by using only references assigned to journals with a single category that is not multidisciplinary. Subsequent iterations include documents classified in previous passes. Milojević concludes that his method is easily replicable, does not require high computational costs, and avoids the problem of labelling categories by using a pre-established classification.

More recently, Álvarez-Llorente et al. (2024) analysed a fractional classification system that uses the ASJC classification scheme from Scopus. This study is based on the multiple generations of references system proposed by Glänzel et al. (2021), but with the possibility of assigning up to five categories instead of three and using different assignment thresholds. This fractional classification system assigns weights to references based on the categories to which the journal containing the reference is assigned in Scopus ("weighted-counting") as opposed to the full-counting system in which all assignments have the same weight. In this study, they compare the results of the three schemes from the model proposed by Glänzel et al. (2021) in their work on multiple generations of references, where they are weighted according to the generation they belong to (denoted as M1, M2, and M3, depending on whether the references are first-generation, second-generation, or both, respectively).

Building on these schemes, new ones are generated where different parameters vary, such as the use of full or fractional counting of weights, the application of averaging or non-averaging

counts to these systems, and the use of three different reference thresholds. For validation, they compare the results with the Scopus fractional assignment scheme (ASJC), with the multidisciplinary area and reallocated miscellaneous categories, as well as with the test collection known as "AAC" (Author's Assignation Collection) (Álvarez-Llorente et al., 2023), where the corresponding authors decide which categories best suit their works. They conclude by stating that higher thresholds for multiple category assignments, along with a system of fractional weights, yield better results. These parameters, using an averaged count with two generations of references, produce what they consider the most appropriate classification for bibliometric use (M3-AWC-0.8).

Even more recently, Álvarez-Llorente et al. (2025) proposed a classification based on references reclassified according to the origin of the citers. This methodology reclassifies Scopus publications, but instead of using the category to which the journals where each article's references are published, are assigned, it uses the category to which the journals of the citing articles belong. Initially, the category with the highest weight is assigned. If the second category exceeds 4/5 of the weight of the first, it is also assigned. Similarly, if the third category exceeds 4/5 of the second, and so on, up to a maximum of 5 categories. In any case, the final weights sum to one and are proportional to the original weights.

With all this in mind, the aim of this study is to determine, through a comparative exercise over a broad period, the variation introduced by the paper level classification system developed by Álvarez-Llorente et al. (2025), described in the previous paragraph, compared to the ASJC-based classification system from the Scopus database. This involves testing various indicators such as normalised impact, excellence distribution, and, in general, analysing how the data varies by examining the areas and categories that constitute the chosen classification. The following research questions are posed:

(a) How does the distribution of documents across areas and categories change with the new classification?
(b) How do documents flow between disciplines with a change in the classification system?
(c) Does the fractional assignment of documents to categories differ between the two systems?
(d) How does normalised impact vary depending on the system used?
(e) Do the excellent documents differ in the new classification system?

2 Materials and methods

The data used in this study come from the Scopus database, specifically from a copy made in March 2024 to which SCImago has access through an agreement with Elsevier. This database employs a classification system known as the All Science Journal Classification (ASJC) (Elsevier, 2023). To enable different researchers to replicate the study, two datasets are provided for online access with the data used (Guerrero-Bote, 2025a,b).

This classification uses a two-level journal-based system, categorising them into 27 thematic areas, which are further subdivided into 311 categories. Among these areas, there is one designated as "*Multidisciplinary*," which is dedicated to purely multidisciplinary journals and does not have any subcategories.

The remaining 26 areas have a specific category dedicated to miscellaneous documents within them, comprising the name of the area followed by the word "*miscellaneous*" in parentheses. For example, within the "*Social Sciences*" area, there is a category named "*Social Sciences (miscellaneous)*".

Álvarez-Llorente et al. (2024) argued that reclassifying works with fewer than three references should be avoided as it would not provide a minimum level of significance.

Due to this, we decided to consider the last 12 years for the study (period 2012–2023), where the values are more stable and there are no significant peaks of difference between them as all values from years prior to 2012 exceed 7% of documents with fewer than three references. As a result, the total number of documents we finally worked with is 35,183,738 items. By working with the total production of documents, we consider that no bias is introduced.

To conduct a coherent comparison, we need to transform the ASJC classification into a fractional classification. For this purpose, we decided to follow the same method presented by Álvarez-Llorente et al. (2023) in which a fractional classification is derived from the ASJC of Scopus. In this regard, the total weight of the assignments for a paper or a journal will sum to one. Therefore, the inverse of the number of ASJC categories assigned is first assigned to each of these assignments.

The scheme used eliminates the general "*Multidisciplinary*" area along with all miscellaneous categories from each area, resulting in a total of 26 thematic areas and 285 categories. However, to retain the values lost by removing the thematic area and miscellaneous categories, a division is performed where "*the weight assigned to an affiliation to the Multidisciplinary subject area is divided among the 285, and the weight assigned to the miscellaneous categories is divided among the rest of the categories of the same subject area*" (Álvarez-Llorente et al., 2023).

With respect to fractional assignment, the objective is to weight the number of assignments a journal has. This method ensures that the value of each assignment varies depending on the number of categories to which the journal is assigned; this value is the weight each assignment receives. The greater the number of assignments, the lower the weight of each individual assignment. Therefore, to determine the size of a specific class, we sum the values of all the weights that comprise it. We will refer to this sum of weights generally as documents.

The fractional classification system based on the ASJC scheme is compared throughout the study with a classification system based on reference reclassification, also of a fractional nature, with the peculiarity that the latter reclassifies references according to the origin of the citers. This system assigns up to a maximum of five categories to each document using a restrictive assignment threshold, thus avoiding excessive multiple category assignments. It is defined in detail in the work by Álvarez-Llorente et al. (2025), and its implementation does not require high computational costs.

Below are some of the bibliometric indicators used in the study:

- Normalised Impact (NI): the average citation normalised received by each document, understood as the ratio between

the citation received by the document and the average citation of documents of the same type, year, and category (Rehn and Kronman, 2008). Since each discipline has different citation habits, the normalised impact indicator is used to make them comparable.

- %Excellence 10 (%Exc10): a document will be considered Excellent10 if it belongs to the top 10% most cited of their same type, year, and thematic area (Bornmann et al., 2012). This indicator is defined as the percentage of "Excellent10" documents. The average of this indicator per scientific category should be 10%, however, there are variations due to the discrete nature of the number of citations.
- %Excellence 1 (%Exc1): the same applies to documents that belongs to the top 1% most cited of their same type, year, and thematic area (Bornmann et al., 2012).

In this section, we also define the concepts of outgoing and incoming documents:

- Outgoing documents: represent the documents (sum of weights) that were previously classified under an area in the ASJC system but are no longer included in that area after changing systems.
- Incoming documents: represent the documents (sum of weights) that, under the new classification system, are assigned to a different area of knowledge than they were in the ASJC classification, thus becoming part of the new area.

3 Results

To present the results in a structured way, we begin by examining the data related to the 26 knowledge areas used by Scopus. Then, we move on to the second part, which focuses on the 285 categories into which the data is organised. Finally, we review the analysis of several bibliometric indicators such as normalised impact and the different types of excellence.

3.1 Areas

Figure 1 compares the documents accumulated by each of the classification systems analysed in this study. The presented graph consists of the 26 knowledge areas used by the Scopus database, each of which is represented by two columns, with each column further divided into two parts.

As defined in the previous section, the sum of weights in the system is considered the number of documents it contains. With this in mind, we define the three variables that make up Figure 1:

- Common System Weight: represented in green, it is the number of documents (sum of weights) assigned to the area by both systems.
- Outgoing Weight: represented in blue, it shows the total number of documents (sum of weights) assigned to the area only by the ASJC system.

- Incoming Weight: represented in orange, it shows the total number of documents (sum of weights) assigned to the area only by the new paper-by-paper system.

Illustrated below is a distribution of the flow network between Scopus areas, showing the flow of documents that occurs when transitioning from the ASJC system to the new classification, developed using the "SCImago Graphica" tool (Hassan-Montero et al., 2022).

This tool allows the representation of networks as shown in Figure 2, where the "LinLog" algorithm (Noack, 2007, 2009) is applied to position the nodes, taking into account the arcs and their weights. Both the nodes and the arcs are scaled according to their weights.

The weight of the nodes is determined by the number of documents assigned to them by the paper-by-paper classification system. Meanwhile, the weight of the links represents the documents that transition from one system to another, specifically from the ASJC classification to the paper level classification.

The clusters identified by the program are generated using the community detection algorithm by Clauset et al. (2004), which distinguishes six main clusters:

- STEM Disciplines (Science, Technology, Engineering and Mathematics).
- Social Sciences.
- Immunology and Microbiology.
- Health Sciences.
- Veterinary.
- Dentistry.

The first cluster is composed of thirteen areas, with the largest being "Engineering", which has a weight of 4,590,331 documents, followed by "Computer Science", with a total of 2,790,049 documents. The strongest flow of information within a single area occurs in this cluster, from "Computer Science" to "Engineering".

The cluster dedicated to the Social Sciences is composed of four main areas, with the largest being "Social Sciences", which has a weight of 2,070,013 documents.

Next is the cluster dedicated to Health Sciences, composed of six nodes, with the largest being "Medicine", which has 8,247,437 documents, making it the largest node in the network. Within this grouping of areas, the most notable link is from the STEM Disciplines cluster, specifically from the area of "Biochemistry, Genetics and Molecular Biology" to "Medicine".

Finally, the clusters dedicated to Veterinary, Immunology, and Dentistry consist of a single area, named "Veterinary", "Immunology and Microbiology" and "Dentistry", respectively.

To highlight the most significant links between the areas, Table 1 lists the links with a weight exceeding 100,000 documents. This table comprises three columns: the first indicates the node from which the link originates, the second indicates the node to which the link leads, and the third shows the strength of the link.

The areas of "Medicine", "Engineering", and "Materials Science" receive the most weight, with the area of "Medicine" receiving the

FIGURE 1
Comparison between the total weights of each Scopus area according to the system used.
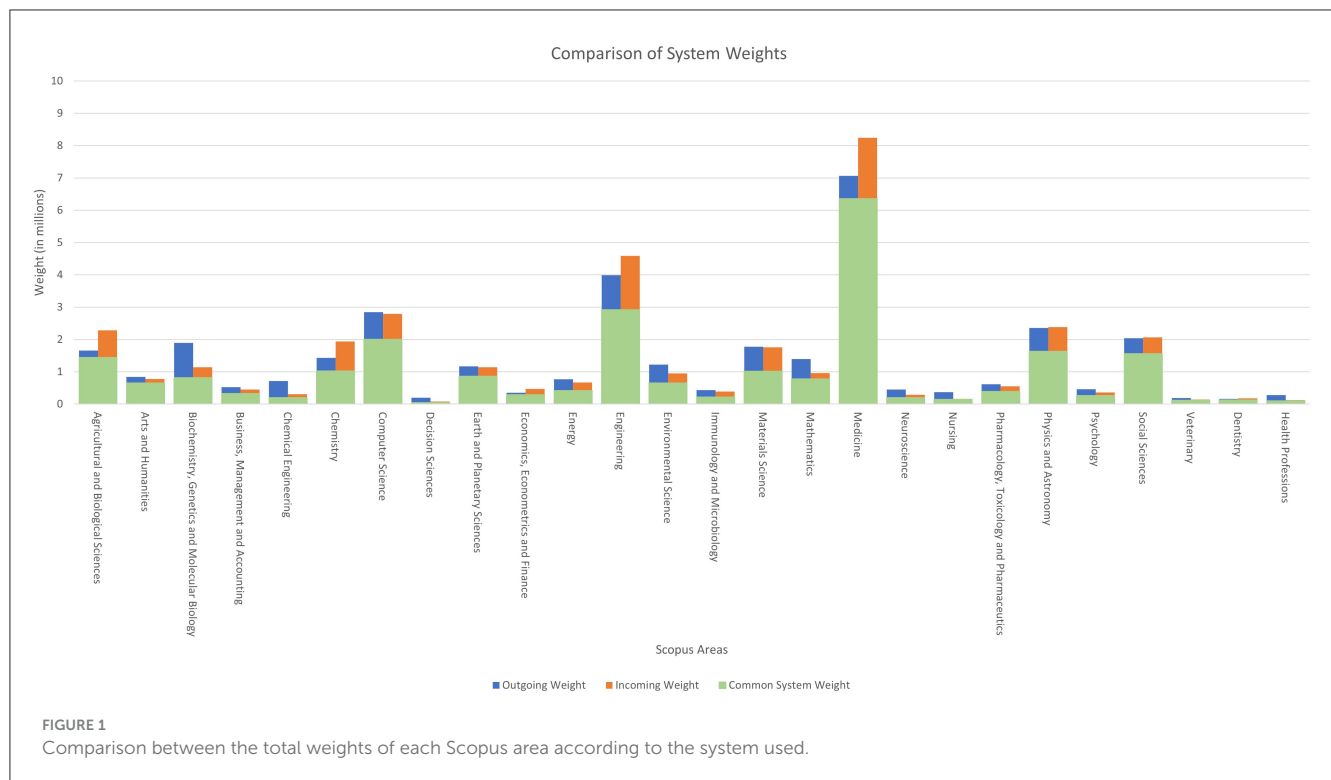
heaviest link from the area of "*Biochemistry, Genetics and Molecular Biology*", with a weight of 554,677 documents.

For the case of second strongest link, it also has a considerable weight transfer, being the one from "*Computer Science*" to "*Engineering*" with 487,984 documents. The link in the opposite direction ranks as the sixth strongest, with a value of 237,567 documents.

Continuing with the study of the thematic areas, a scatter plot is included, showing the Scopus areas according to the percentage of incoming or outgoing documents they have with respect to the paper-by-paper system (Figure 3). In this representation, the nodes with a higher percentage of incoming weight are positioned to the right, while the areas with a higher percentage of outgoing weight are placed at the top.

Since we are comparing the outgoing and incoming documents with respect to the paper-by-paper system, some of the percentages shown in the diagram exceed 100%. This is because the number of documents that move from one area to another under the new system is greater than the number of documents that remain in the same area under the new system, thus exceeding 100%. This occurs only in four areas of knowledge: Nursing, Health Professions, Chemical Engineering, and Decision Sciences.

We highlight a section of the graph where more than half of the areas are grouped; this occurs between the ranges of 20%–40% incoming weight and 0%–60% outgoing weight, encompassing a total of 14 areas.

To conclude with the data related to the areas, the average weight of each Scopus area according to the classification system used is analysed, along with the percentage of documents that belong exclusively to that category. Figure 4 uses two different axes of representation:

- The left axis ranges from 0% to 100% and is represented by blue and orange columns. The blue column indicates the percentage of documents assigned to a single category in the ASJC system, while the orange column represents this percentage in the paper level system.
- The right axis ranges from 0 to 1 and is represented in the figure by blue circles and orange triangles. This axis represents the average weight of each document within the area.

Starting with the columns, it is observed that the percentage values of assignment to an area in the ASJC classification are lower than those in the new classification. These values range from 5.61% to 66.49% for the ASJC classification, whereas for paper level system they range from 57.88% to 90.90%.

In the case of average weights per document, it is also notable that the values are higher with the new system. For the ASJC classification system, the area of "*Medicine*" stands out with an associated average weight of 0.65, followed by the area of "*Social Sciences*" with a value of 0.49 and the areas of "*Agricultural and Biological Sciences*" and "*Computer Science*" with a value of 0.48.

Similar to the journal-based classification, the paper-by-paper system also shows that the area with the highest average weight per document is "*Medicine*" with a value of 0.92, followed by the areas of "*Agricultural and Biological Sciences*" and "*Computer Science*" with values of 0.89 and 0.85, respectively.

Due to the weights of each document being less fractionated in the paper level classification compared to the ASJC classification, the percentage of documents assigned to a single category is higher in the new system. It can be observed that the average weights in the new classification system are consistently higher than those in the old system.

FIGURE 2
Scopus areas network diagram.
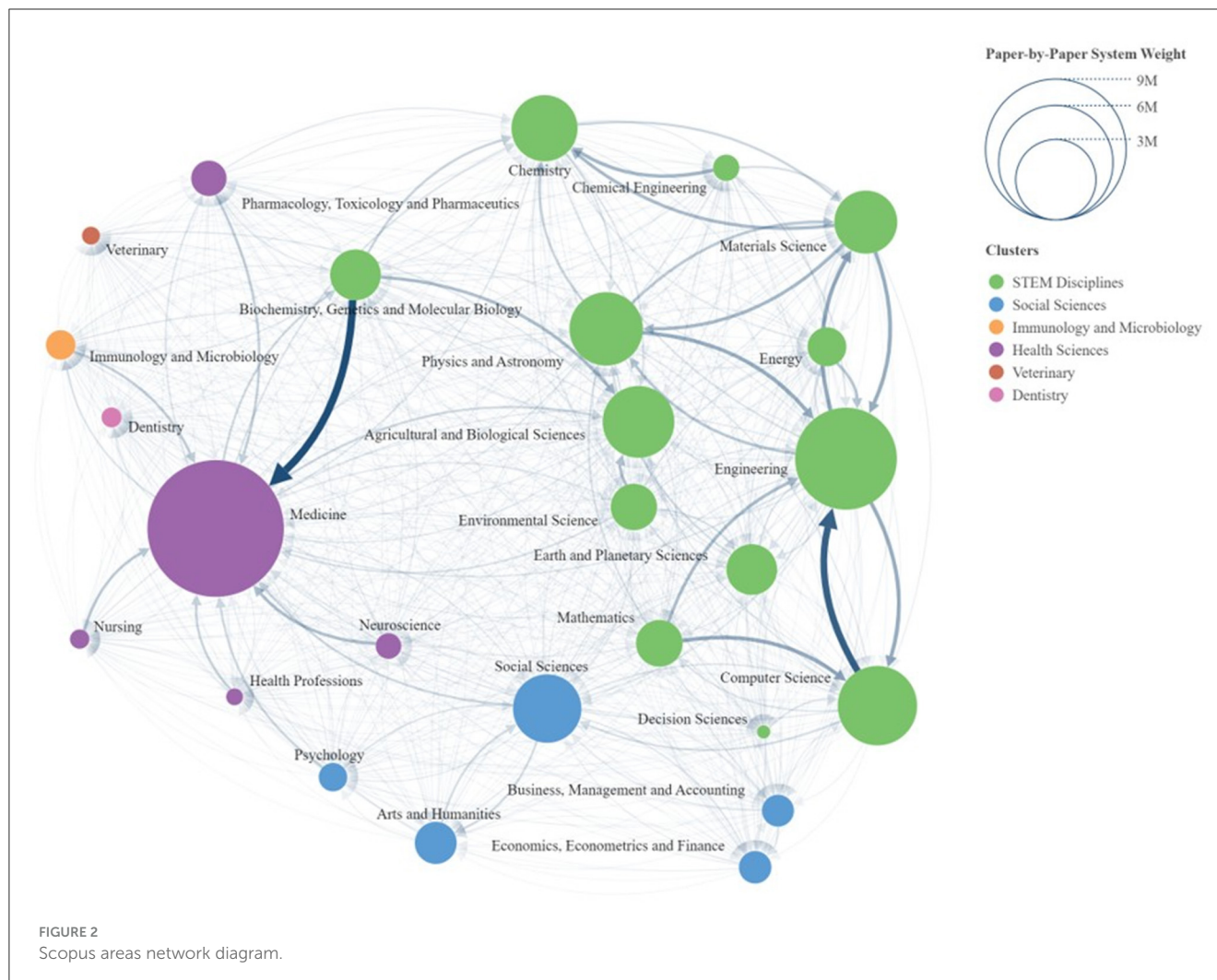
3.2 Categories

Next, a bar chart is included to visually represent the distribution of categories by their weight according to the system used. Figure 5 shows this diagram, which is divided into intervals of 100,000 documents. The blue columns represent the ASJC classification, while the orange bars represent the paper level system.

This diagram, functioning as a histogram, illustrates how many of the Scopus categories fall within each interval according to the number of documents they contain.

The two systems share similarities, highlighting how in both cases more than 50% of the categories are grouped within the first interval. For the ASJC classification, 52.28% of the categories consist of fewer than 100,000 documents, with the next bracket containing 31.58% of the categories in the 100,000–200,000 document range. With the new classification system, nearly 70% of the categories are composed of fewer than 100,000 documents, with a total of 69.47%.

The new classification system has a peculiarity when dealing with categories that are at the extremes of size: it tends to decrease the smaller categories and further increase the larger ones. This

is why the percentage of categories with fewer than 100,000 documents are almost 17% higher in the new classification system compared to the ASJC.

In both cases, the representations have a negative trend, starting with the highest point in the first interval and progressively decreasing. There are some intervals that do not contain any categories, these are the ones that have a value of zero.

It is important to highlight that there is a specific category in the paper-by-paper classification system that stands out due to its highest value. The category in question is "*Electrical and Electronic Engineering*," which has a total weight of 2,027,561 documents. The next category is "*Oncology*", with a total of 1,022,225 documents.

To show the general trend of the categories concerning information flows, the results are grouped in a table (Table 2). The first part is dedicated to the weights of the categories according to the classification system used (titled "Classification System"), analysing the average and coefficient of variation (Cramer and Howitt, 2004) (multiplied by 100 to express it as a percentage). The second part (titled "Outgoing Flow") deals with the outgoing flows with respect to the paper level classification. Finally (titled "Incoming Flow") we perform the same analysis but for the incoming flows.

TABLE 1 Stronger links between Scopus areas.

Outgoing flow	Incoming flow	Number of documents
Biochemistry, Genetics and Molecular Biology	Medicine	554,677
Computer Science	Engineering	487,984
Engineering	Materials Science	251,137
Mathematics	Computer Science	244,036
Materials Science	Engineering	237,686
Engineering	Computer Science	237,567
Physics and Astronomy	Engineering	237,242
Chemical Engineering	Chemistry	207,289
Materials Science	Physics and Astronomy	202,166
Materials Science	Chemistry	190,198
Mathematics	Engineering	186,563
Biochemistry, Genetics and Molecular Biology	Agricultural and Biological Sciences	182,003
Neuroscience	Medicine	176,740
Environmental Science	Agricultural and Biological Sciences	169,053
Energy	Engineering	156,043
Nursing	Medicine	150,441
Engineering	Physics and Astronomy	149,653
Physics and Astronomy	Materials Science	148,419
Immunology and Microbiology	Medicine	132,625
Pharmacology, Toxicology and Pharmaceutics	Medicine	124,767
Health Professions	Medicine	117,915
Physics and Astronomy	Chemistry	117,182
Biochemistry, Genetics and Molecular Biology	Chemistry	116,030
Chemistry	Materials Science	111,797
Medicine	Biochemistry, Genetics and Molecular Biology	102,320

Starting with the differences between the classification systems, the first notable point is that, although the average number of documents in both systems is the same, the coefficient of variation is not. Documents in the ASJC classification system are distributed among the categories more homogeneously, with a variation of 93.31% compared to the 163.51% variation in the paper level system.

For the cases of outgoing and incoming flows, a similar pattern is observed. Despite the average of both variables being 67,424 documents, the coefficient of variation is higher for incoming flows than for outgoing flows. Given these high variation values, it can be stated that the new system increases the size of the larger categories and reduces the size of the smaller categories.

Both the average and the coefficient of variation for outgoing flow are higher than for incoming flow. The average value for incoming flow is 37.26%, while for outgoing flow it reaches 253.23%. Additionally, the incoming flow is more homogeneous, with a value of 52.11% compared to the 150.28% for outgoing flow. This is because, when working with percentage values, the difference is much more pronounced when there is an outgoing flow from smaller categories compared to when the incoming values are directed towards larger categories, resulting in a smaller percentage difference.

Similar to Table 1, another table (Table 3) is designed to show the significant flows between categories, including all those that exceed 40,000 documents.

The categories of "Oncology" and "Electrical and Electronic Engineering" receive the most weight, with the field of oncology occupying the first position and the next four positions held by the latter category. The heaviest link is from "Cancer Research" to "Oncology", with a weight of 135,709 documents. The areas associated with these fields of knowledge correspond to the areas linked by the heaviest connexion in Table 1.

The top-ranked link has a noteworthy peculiarity: the incoming and outgoing categories belong to different communities according to the algorithm used in Table 3. The outgoing category is part of the "Biochemistry, Genetics and Molecular Biology" area, which is included in the STEM disciplines cluster, while the incoming category falls within the "Medicine" area, in the Health Sciences cluster.
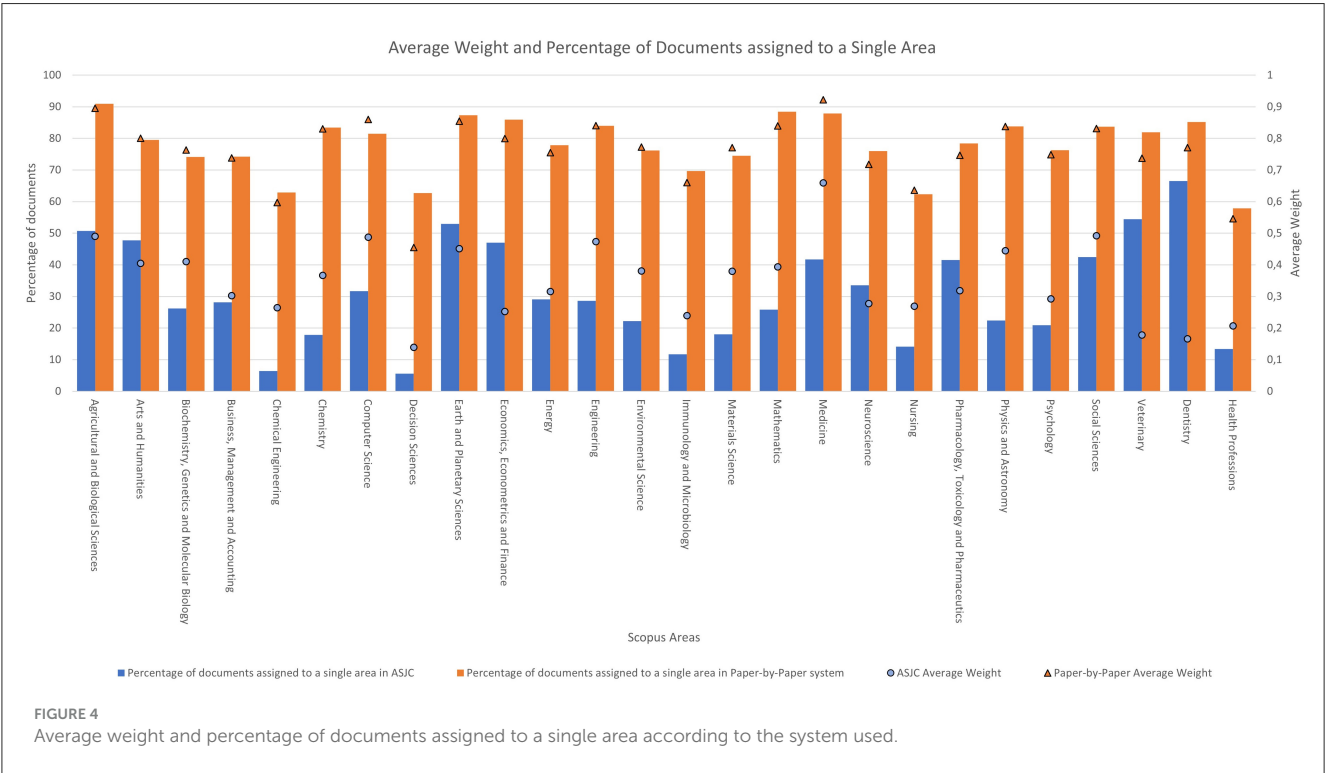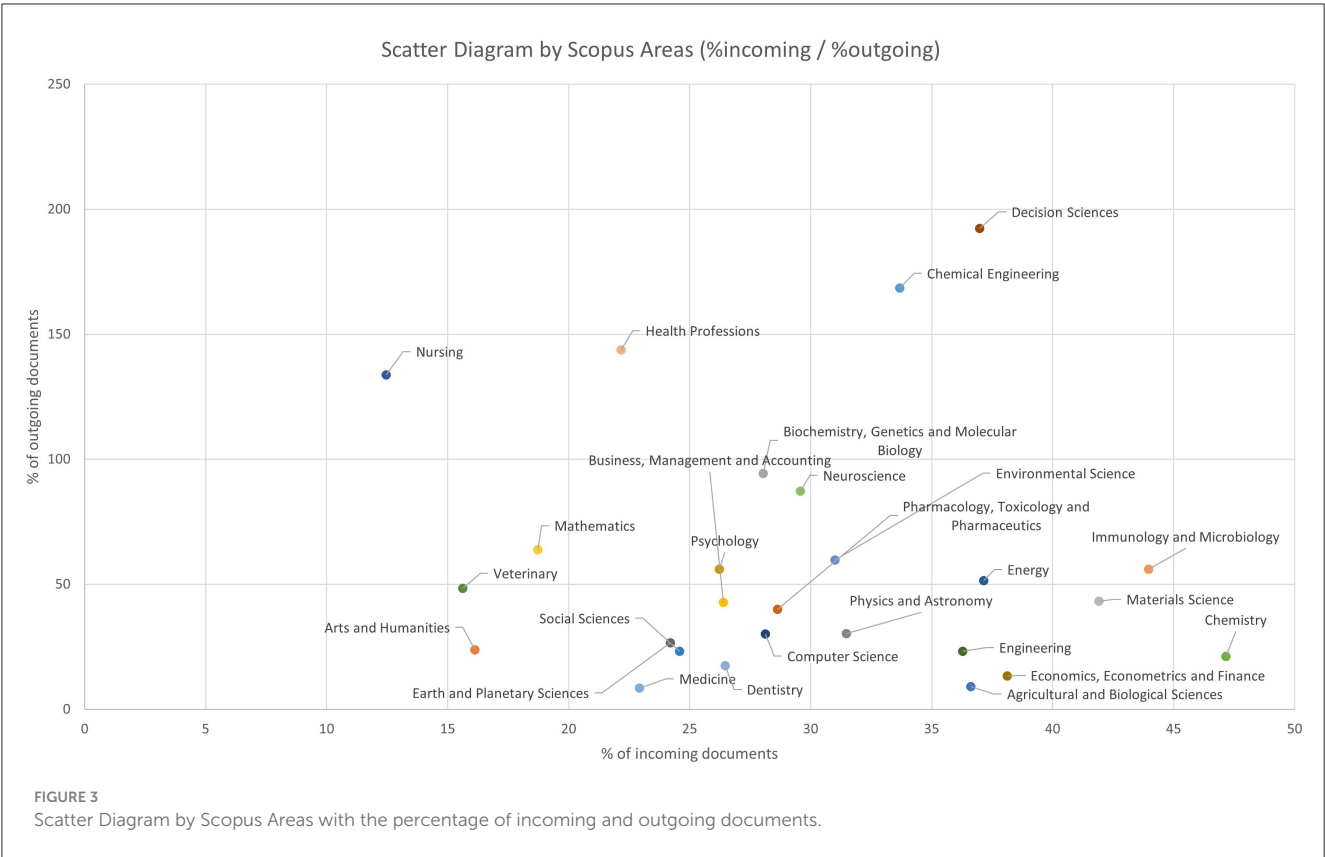
Finally, Table 4 is included to explain the differences in category weights according to the classification system. It is divided into two parts:

- The first part is dedicated to the ASJC system, which includes the average weight of Scopus categories and their average percentage of assignment to a single category, along with their respective coefficients of variation.
- The second part is dedicated to the paper-by-paper system, where the structure is similar but this time with the data from the new system.

Focusing on the variables of the assigned weights, it is evident that the value of the categories is, on average, lower in the ASJC system, with an average of 0.059, meaning that each document is assigned to ∼17 categories on average, compared to 0.324, three categories, in the paper-by-paper system.

Despite both systems registering a very similar coefficient of variation for this variable, the weight assigned to each category is higher on average in the paper-by-paper system, resulting in less fractional distribution of weights among the different categories.

Regarding the variable that measures the percentage of documents assigned to a single category, the superiority of the paper-level system over the ASJC is notable, with more than half of the documents being assigned to a single category compared to 13.51% in the ASJC classification. Additionally, the coefficient of variation in the ASJC system is 96.74%, while in the paper-level system it is 44.69%, providing not only higher average values but also more homogeneous ones.

FIGURE 3
Scatter Diagram by Scopus Areas with the percentage of incoming and outgoing documents.

FIGURE 4
Average weight and percentage of documents assigned to a single area according to the system used.

3.3 Bibliometrics indicators

To assess how the change in classification affects quality indicators, we compare the values provided by the normalised impact quality indicator and the excellence indicator.

Figure 6 represents the annual average evolution of the absolute value of the difference in each paper between the normalised impact calculated from the two classifications.

Since citation data are not comparable across disciplines due to variations between them, it is important that the normalised
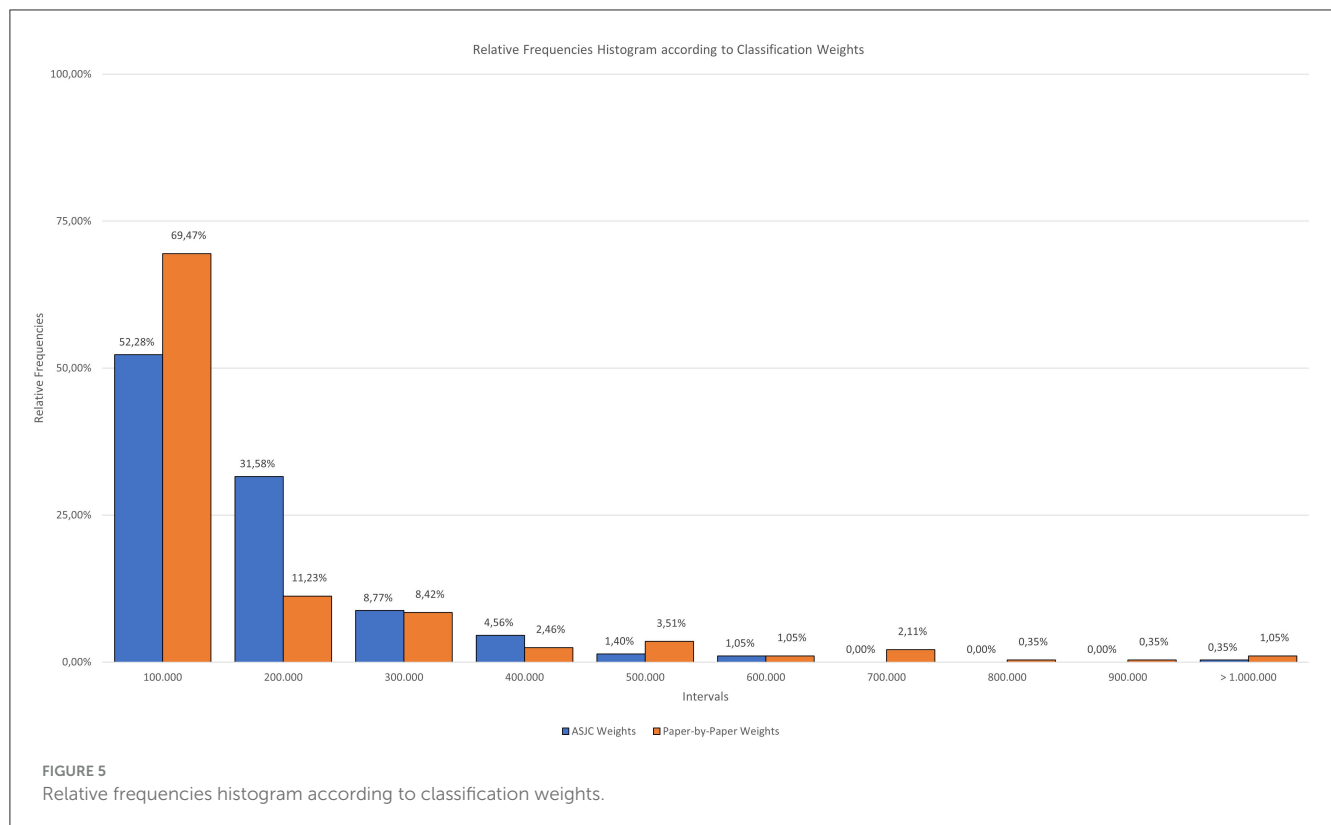
FIGURE 5
Relative frequencies histogram according to classification weights.

TABLE 2 Summary measures of incoming and outgoing flow.

	Average	Coefficient of variation
Classification system		
Weight in the ASJC system	123,451	93.31%
Weight in the paper-by-paper system	123,451	163,51%
Outgoing flow		
Outgoing flow	67,424	81.40%
Percentage of outgoing flow with respect to paper-by-paper	253.23%	150.28%
Incoming flow		
Incoming flow	67,424	191.85%
Percentage of incoming flow with respect to paper-by-paper	37.26%	52.11%

impact indicator tries to homogenise the citation habits of similar documents.

The paper level system shows a 22% variation in the average normalised impact compared to the traditional ASJC system, indicating that the normalised impact varies significantly depending on the classification used.

Although the rest of the figures extend the study up to the year 2023, this year is excluded from this figure because it has had a shorter citation window, making the values less precise than those of the previous years.

We also include a table (Table 5) that compares the annual differences in normalised impact between both systems. It consists of four columns:

- Year.
- Median and Standard Deviation in the ASJC system.
- Median and Standard Deviation in the Paper-by-Paper system.
- Correlation: it is the correlation between the normalised impact of each paper in one system and the other.

We can highlight how the distribution maintained by both systems is very similar, as both the median values and the standard deviation they offer are different but very similar. Additionally, the correlation data value is, on average, 0.93, which means that when a paper has a high impact in one system, it tends to have a high impact in the other as well. The differences and correlations are statistically significant due to the large sample size.

Figure 7 shows a comparison by areas, represented by a column chart that illustrates the standard deviation of the normalised impact according to the classification used. The variable "*STD NI ASJC*" is represented in blue, indicating the standard deviation for the ASJC classification, while the variable "*STD NI Paper-by-Paper*" is represented in orange, indicating the standard deviation for the paper level classification.

In the ASJC classification, high values are notable in areas such as "*Computer Science*", "*Arts and Humanities*", and "*Mathematics*". Meanwhile, in the paper-by-paper classification, the areas with the highest standard deviation in normalised impact are "*Computer Science*", "*Business, Management and Accounting*" and "*Mathematics.*"

TABLE 3 Stronger links between Scopus categories.

Outgoing flow	Incoming flow	Number of documents
Cancer Research	Oncology	135,709
Energy Engineering and Power Technology	Electrical and Electronic Engineering	79,875
Computer Science Applications	Electrical and Electronic Engineering	77,289
Electronic, Optical and Magnetic Materials	Electrical and Electronic Engineering	72,584
Control and Systems Engineering	Electrical and Electronic Engineering	69,991
Computer Networks and Communications	Electrical and Electronic Engineering	59,130
Space and Planetary Science	Astronomy and Astrophysics	59,085
Electronic, Optical and Magnetic Materials	Condensed Matter Physics	58,381
Finance	Economics and Econometrics	55,233
Instrumentation	Electrical and Electronic Engineering	53,710
Mechanics of Materials	Mechanical Engineering	53,101
Instrumentation	Atomic and Molecular Physics, and Optics	49,930
Electronic, Optical and Magnetic Materials	Atomic and Molecular Physics, and Optics	49,286
Neurology	Neurology (clinical)	49,241
Condensed Matter Physics	Electrical and Electronic Engineering	48,398
Theoretical Computer Science	Software	47,388
Building and Construction	Civil and Structural Engineering	46,408
Hardware and Architecture	Computer Networks and Communications	44,986
Information Systems	Computer Networks and Communications	44,957
Atomic and Molecular Physics, and Optics	Condensed Matter Physics	43,897
Signal Processing	Electrical and Electronic Engineering	41,395
Physical and Theoretical Chemistry	Organic Chemistry	40,959
Computer Science Applications	Computer Networks and Communications	40,602
Hardware and Architecture	Electrical and Electronic Engineering	40,378

TABLE 4 Summary measures of categories weights.

System	Category	Average	Coefficient of variation
ASJC System	Weight assigned to categories	0.059	69.22%
	Percentage belonging to a single category	13.51%	96.74%
Paper-by-Paper System	Weight assigned to categories	0.324	66.37%
	Percentage belonging to a single category	56.08%	44.69%

The new system outperforms ASJC in only 76 out of 285 areas, with a very similar difference in both cases.

The results show that the paper level classification tends to offer lower standard deviation values compared to the ASJC classification, indicating that it is more homogeneous in terms of normalised impact.

For the excellence indicators, we work with the indicators described in the methodology: *%Exc10* and its variants in Figure 8, and *%Exc1* and its variants in Figure 9.

Starting with Figure 8, consists of three columns: the first, in green, labelled as *%Exc10*, represents the documents that are excellent in both classifications systems. In orange, we define "*%Exc10 non-ASJC*" as the percentage of documents that were not considered excellent in the ASJC classification but are now considered excellent with the paper-by-paper classification. Conversely, "*%Exc10 ASJC*" is defined in blue as the percentage of documents that were considered excellent in the ASJC classification but are not considered excellent in the paper-by-paper classification.

It is evident that the categories most benefited by the new classification are "*Arts and Humanities*", "*Nursing*", "*Veterinary*", and "*Health Professions*," showing the most notable differences between their "*Exc10 non-ASJC*" and "*Exc10 ASJC*" values.

There is a relationship between the areas with a smaller number of documents in the top 10% of excellence and those that achieve higher values in the paper level classification. If the areas are ordered from lowest to highest by their "*%Exc10*", we find that the same areas that benefited from the classification change, mentioned in the previous paragraph, are the ones with the lowest "*%Exc10*" values.

In other words, the areas with a lower percentage of documents belonging to the top 10% of their area correspond to the areas with the highest percentage in the indicator of documents that become excellent in paper level but were not considered excellent in ASJC.

For the case of 1% excellence, the process is similar to the previous one, but focusing on the top 1% of works, as shown in Figure 9. With this indicator, the four areas with the lowest "*%Exc1*" are again the four previously mentioned, which then benefit the most in "*%Exc1 non-ASJC*".

Although the "*Arts and Humanities*" category achieves very high excellence values in the new classification for both types of
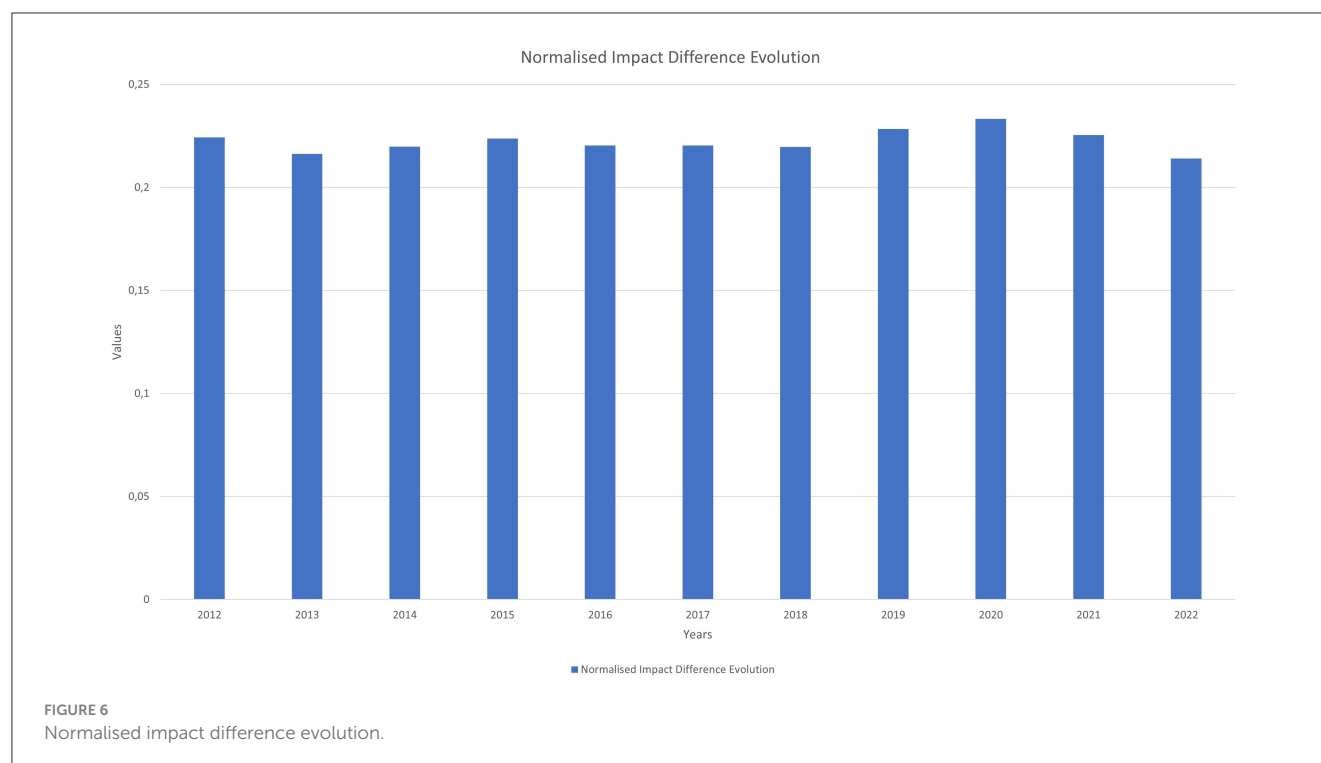
On average, the ASJC system has a standard deviation of normalised impact that is 0.53 higher. The new system outperforms ASJC in only 3 out of 26 areas, with a very small difference. When performing the same analysis at the category level, the ASJC system has a Standard Deviation of normalised impact that is 0.44 higher.

FIGURE 6
Normalised impact difference evolution.

TABLE 5 Differences in normalised impact between systems.

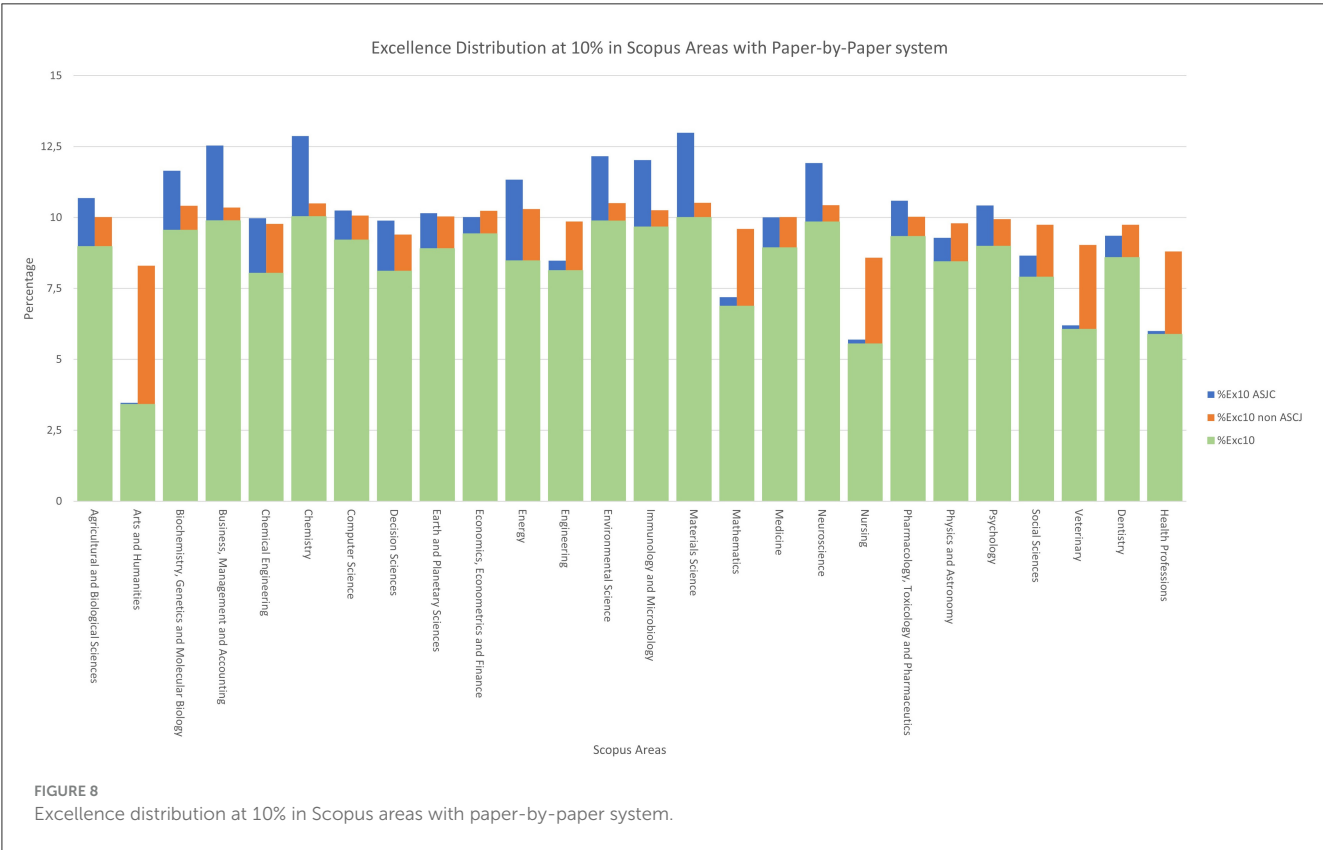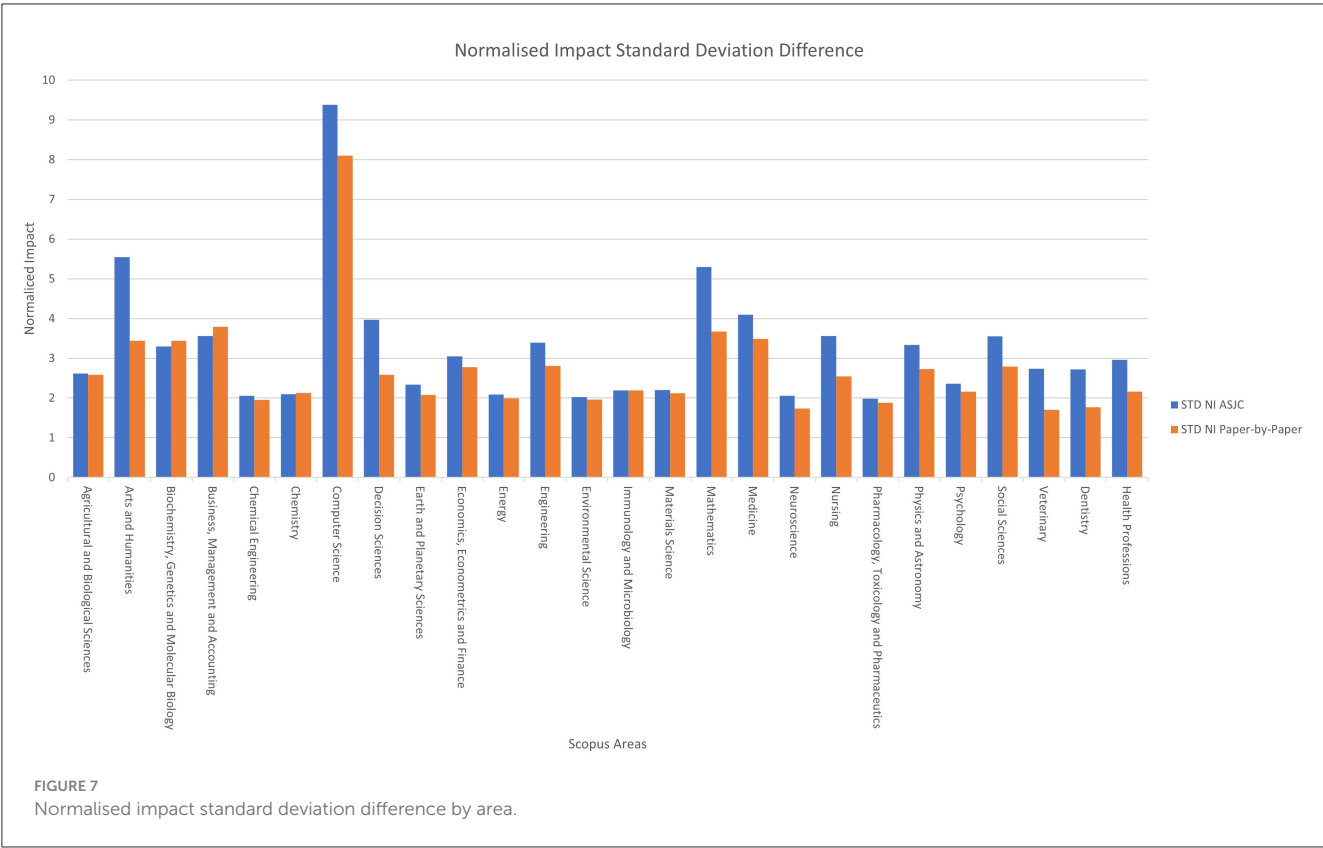| Year	ASJC		Paper-by paper system		Correlation
	Median	Standard deviation	Median	Standard deviation	
2012	0.37	5.83	0.38	5.12	0.97
2013	0.39	3.77	0.40	3.45	0.92
2014	0.40	4.49	0.41	3.76	0.92
2015	0.42	5.68	0.42	4.26	0.89
2016	0.43	5.20	0.44	4.30	0.94
2017	0.45	4.03	0.45	3.50	0.93
2018	0.46	2.84	0.46	2.72	0.92
2019	0.47	3.30	0.48	3.21	0.95
2020	0.46	3.55	0.47	2.80	0.90
2021	0.48	3.88	0.49	3.41	0.93
2022	0.43	2.52	0.44	2.44	0.94

excellence, this is directly related to the way bibliometric indicators are measured in these disciplines. The dissemination of knowledge through articles is not as frequent as in purely scientific disciplines, so these data should be treated separately.
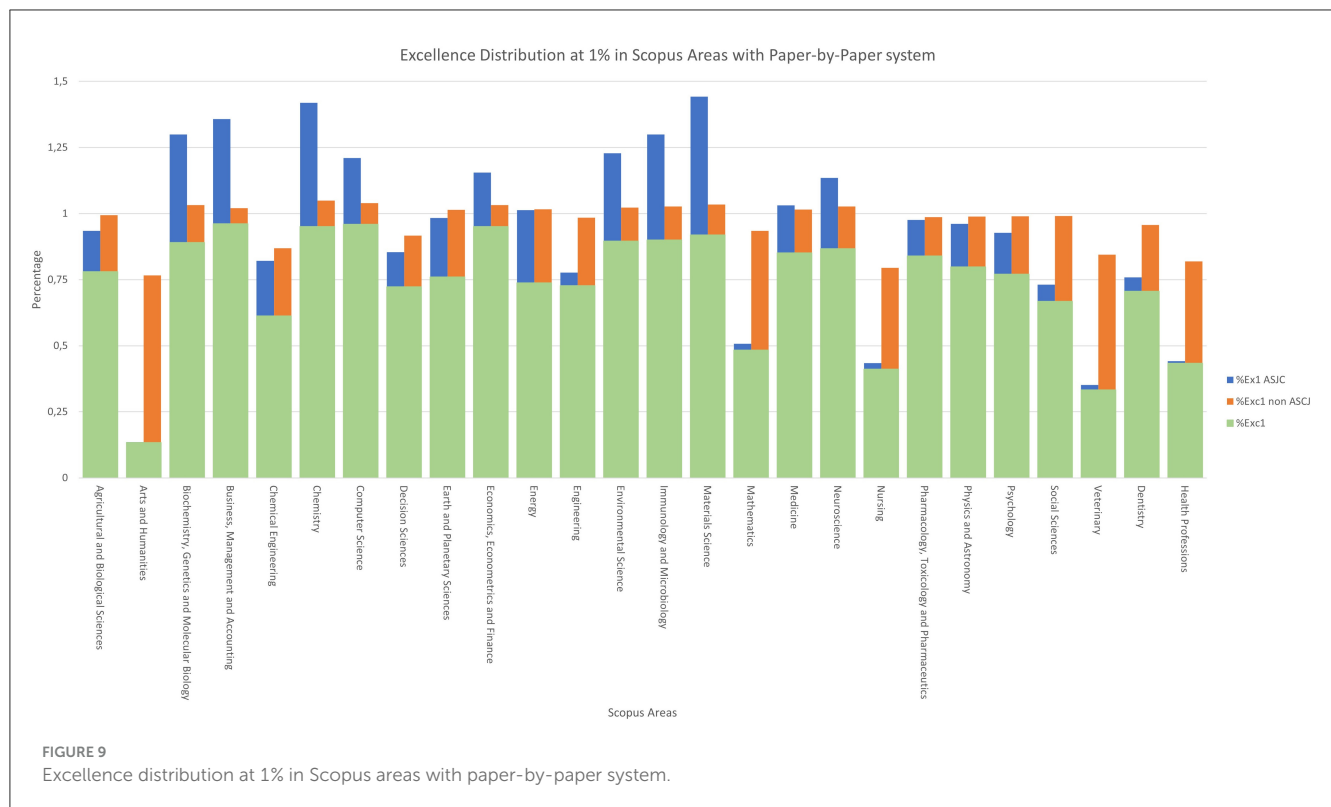
4 Conclusions

After conducting a comparative analysis between the two classification systems, several conclusions can be drawn based on the different aspects addressed.

Starting with the limitations, the new system is less reliable in classifying papers with few references, and there are still papers without references in the database. On the other hand, regarding the distribution of documents across disciplines, the paper-by-paper system, far from balancing the size of the ASJC areas and categories, tends to accentuate differences at the extremes, increasing the size of both the largest areas and categories while decreasing the smallest ones.

Logically, the areas composed of a larger number of documents are the ones that incorporate (or lose) the most documents, as seen in the graphical representation in Figure 2, where the areas

FIGURE 7
Normalised impact standard deviation difference by area.

FIGURE 8
Excellence distribution at 10% in Scopus areas with paper-by-paper system.

FIGURE 9
Excellence distribution at 1% in Scopus areas with paper-by-paper system.

of "*Medicine*" and "*Engineering*" are the largest. In percentage terms, areas such as "*Decision Sciences*" or "*Health Professions*", which are the smallest in size, have very high percentages of outgoing documents. Similarly, the areas of "*Chemical Engineering*" and "*Nursing*" exhibit this trend but with lower values. These are two fields of knowledge that are among the smallest in terms of size, but the percentage of outgoing documents exceeds 100%.

Another area worth highlighting, in this case for the percentage of incoming documents, is "*Immunology and Microbiology*". Despite being the ninth smallest area, it achieves the second-highest percentage of incoming documents, with nearly 45%.

This characteristic of increasing size disparities is not desirable for scientometrics, but it may highlight the lesser significance of some categories, where very few works cite the intellectual foundations specific to them.

The new system, compared to ASJC, reduces the number of fractions into which each document is divided. This characteristic aligns with the findings of Álvarez-Llorente: "*the assignment of a reduced number of categories per paper is a highly desirable feature*" (Álvarez-Llorente et al., 2025).

Since the weights are less fractionated in the new system, the values for assignment to a single area in the ASJC classification have an average of 30%, while with the paper level system, this value increases to nearly 78% of the documents. In all areas, a higher percentage of documents are assigned to a single category in the new system.

This is indeed a desirable characteristic because it better represents reality. As previously indicated, there are multidisciplinary journals that publish works from various disciplines, but far fewer works are genuinely multidisciplinary or interdisciplinary.

For the case of normalised impacts, there is a significant variation between them depending on the classification used. As shown in Figure 6, the value resulting from calculating the variable using the annual average evolution of the absolute value of the difference between the normalised impact of the paper-by-paper system and the ASJC system indicates that, on average, this figure varies by 22%.

This new classification offers a more homogeneous distribution of normalised impacts, grouping documents with similar citation habits together. The standard deviation values shown by the ASJC classification system average 3.24, compared to 2.71 in the new system. This characteristic is also desirable because one of the main uses of scientometric classifications is to compare works belonging to disciplines with the same citation habits, and this indicates that in the new system, the citation habits within each category are more homogeneous.

Regarding the issues related to excellence indicators, there is a common percentage of excellent documents that remain in both classifications: for documents belonging to the top 10% of excellence, this figure is 8.72%, while for the 1% excellence indicator, this value is 0.79% of documents in common.

Finally, we can state that the number of excellent documents in each area varies depending on the classification used. The new system ensures that the areas that do not reach the 10% or 1% excellence thresholds get closer to the respective values of 10 and 1. Logically, this characteristic of the new system is also desirable.

We consider that the use of this paper-by-paper system offers a series of advancements with the potential to improve bibliometric

systems. We particularly highlight an improvement in accuracy compared to journal-based systems when classifying scientific documents. This represents a substantial improvement in the field of bibliometrics, benefiting researchers, publishers, journals, and institutions alike.

In conclusion, we believe that the new system, overall, represents in a better way the reality and exhibits more desirable characteristics.

5 Future research

As future lines of research, we propose to investigate how this new system might affect the indicators of countries and institutions.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found: Guerrero-Bote, V. P. (2025a). Classification for Scopus 2024 based on references reclassified by the origin of the papers citing them (2012-2022) (Versión V.1) [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.14713511. Guerrero-Bote, V. P. (2025b). Classification for Scopus 2024 journal-based adapted to a fractional mode (2012-2022) (Versión V.1) [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.14713710.

Author contributions

MP-R: Formal analysis, Visualization, Writing – original draft. RG-C: Conceptualization, Formal analysis, Methodology, Validation, Visualization, Writing – review & editing. VG-B: Conceptualization, Data curation, Funding acquisition, Methodology, Software, Supervision, Writing – review & editing. FM-A: Conceptualization, Funding acquisition, Resources, Validation, Writing – review & editing.

Conflict of interest

FM-A was employed at SCImago Research Group.

The remaining authors declare that that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Álvarez-Llorente, J. M., Guerrero-Bote, V. P., and de Moya-Anegón, F. (2024). *New fractional classifications of papers based on two generations of references and on the ASJC scopus scheme. Scientometrics* 129, 3493–3515. doi: 10.1007/s11192-024-05030-2

Álvarez-Llorente, J. M., Guerrero-Bote, V. P., and de Moya-Anegón, F. (2025). New paper-by-paper classification for Scopus based on references reclassified by the origin of the papers citing them. *J. Informetr.* 6:101647. doi: 10.1016/j.joi.2025.101647

Álvarez-Llorente, J. M., Guerrero-Bote, V. P., and De-Moya-Anegón, F. (2023). Creating a collection of publications categorized by their research guarantors into the Scopus ASJC scheme. *Profesional De La Información* 32. doi: 10.3145/epi.2023.dic.04

Bornmann, L., and Daniel, H. (2008). What do citation counts measure? A review of studies on citing behavior. *J. Document.* 64, 45–80. doi: 10.1108/00220410810844150

Bornmann, L., Moya-Anegon, F., and Leydesdorff, L. (2012). The new Excellence Indicator in the World Report of the SCImago Institutions Rankings 2011. *J. Informetr.* 6, 333–335. doi: 10.1016/j.joi.2011.11.006

Boyack, K. W., and Klavans, R. (2010). Co-citation analysis, bibliographic coupling, and direct citation: which citation approach represents the research front most accurately? *J. Am. Soc. Inform. Sci. Technol.* 61, 2389–2404. doi: 10.1002/asi.21419

Boyack, K. W., Newman, D., Duhon, R. J., Klavans, R, Patek, M., Biberstine, J. R., et al. (2011). Clustering more than two million biomedical publications:

comparing the accuracies of nine text-based similarity approaches. *PLoS ONE* 6:e18029. doi: 10.1371/journal.pone.0018029

Boyack, K. W., Small, H., and Klavans, R. (2013). Improving the accuracy of co-citation clustering using full text. *J. Am. Soc. Inform. Sci. Technol.*, 64, 1759–1767. doi: 10.1002/asi.22896

Clauset, A., Newman, M. E. J., and Moore, C. (2004). Finding community structure in very large networks. *Phys. Rev. E.* 70:066111. doi: 10.1103/PhysRevE.70.066111

Cramer, D., and Howitt, D. (2004). "Coefficient of variation," in *The SAGE Dictionary of Statistics*. Available at: https://methods.sagepub.com/dict/edvol/the-sage-dictionary-of-statistics/chpt/coefficient-variation

Elsevier (2023). *What is the Complete List of Scopus Subject Areas and All Science Journal Classification Codes (ASJC)?* Available at: https://service.elsevier.com/app/answers/detail/a_id/15181/supporthub/scopus (accessed August 15, 2024)

Glänzel, W., and Schubert, A. (2003). A new classification scheme of science fields and subfields designed for scientometric evaluation purposes. *Scientometrics* 56, 357–367. doi: 10.1023/A:1022378804087

Glänzel, W., Schubert, A., and Czerwon, H-. J. (1999a). An item-by-item subject classification of papers published in multidisciplinary and general journals using reference analysis. *Scientometrics* 44, 427–439. doi: 10.1007/BF02458488

Glänzel, W., Schubert, A., Schoepflin, U., and Czerwon, H. J. (1999b). An item-by-item subject classification of papers published in journals covered by the SSCI database using reference analysis. *Scientometrics* 46, 431–441. doi: 10.1007/BF0245 9602

Glänzel, W., Thijs, B., and Huang, Y. (2021). "Improving the precision of subject assignment for disparity measurement in studies of interdisciplinary research," in *En FEB Research Report MSI_2104* (Kortrijk: FEB - KU Leuven). pp. 1–12.

Guerrero-Bote, V. P. (2025a). *Classification for Scopus 2024 Based on References Reclassified by the Origin of the Papers Citing Them (2012–2022) (Versión V.1) [Dataset]*. Genève: Zenodo.

Guerrero-Bote, V. P. (2025b). *Classification for Scopus 2024 Journal-Based Adapted to a Fractional Mode (2012–2022) (Versión V.1) [Dataset]*. Genève: Zenodo.

Hassan-Montero, Y., De-Moya-Anegón, F., and Guerrero-Bote, V. P. (2022). SCImago Graphica: a new tool for exploring and visually communicating data. *Profesional De La información* 31. doi: 10.3145/epi.2022.sep.02

Janssens, F., Glänzel, W., and De Moor, B. (2008). A hybrid mapping of information science. *Scientometrics* 75, 607–631. doi: 10.1007/s11192-007-2002-7

Janssens, F., Zhang, L., Moor, B. D., and Glänzel, W. (2009). Hybrid clustering for validation and improvement of subject-classification schemes. *Inform. Process. Manage.* 45, 683–702. doi: 10.1016/j.ipm.2009.06.003

Lin, J., and Wilbur, W. J. (2007). PubMed related articles: a probabilistic topic-based model for content similarity. *BMC Bioinform.* 8:423. doi: 10.1186/1471-2105-8-423

Milojević, S. (2020). Practical method to reclassify Web of Science articles into unique categories and broad disciplines. *Q. Sci. Stud.* 1, 183–206. doi: 10.1162/qss_a_00014

Noack, A. (2007). Energy models for graph clustering. *J. Graph Algorithms Appl.* 11, 453–480. doi: 10.7155/jgaa.00154

Noack, A. (2009). Modularity clustering is force-directed layout. *Phys. Rev. E Stat. Nonlinear Soft Matter Phys.* 79:026102. doi: 10.1103/PhysRevE.79.026102

Rehn, C., and Kronman, U. (2008). *Bibliometric Handbook for Karolinska Institutet*. Huddinge: Karolinska Institutet.

Waltman, L., and van Eck, N. J. (2012). A new methodology for constructing a publication-level classification system of science. *J. Am. Soc. Inform. Sci. Technol.* 63, 2378–2392. doi: 10.1002/asi.22748