Fraunhofer ISI Discussion Papers *Innovation Systems and Policy Analysis* No. 33
ISSN 1612-1430
Karlsruhe, February 2013

Systematic Analysis of Coverage and Usage of Conference Proceedings in Web of Science

Subproject in the Kompetenzzentrum Bibliometrie

Carolin Michels
Jun-Ying Fu

Fraunhofer Institute for Systems
and Innovation Research ISI

Contents Page

1 Introduction/Aim of this study 1

2 Methodology 3

2.1 State of the Art 3

2.2 Proposed Approach 6

3 Exemplary comparison of bibliometric indicators for selected countries 9

4 Results 17

4.1 Descriptive Analysis 17

4.2 Implications for Bibliometric Analyses 37

5 Conclusion 52

References 53

Annex 56

Tables

Table 1: Rank of publication shares for journal publications and conference proceedings for selected countries in 2009 10

Table 2: Rank of citation rate for journal publications and conference proceedings for selected countries in 2007 12

Table 3: Rank of IA values for journal publications and conference proceedings for selected countries in 2007 14

Table 4: Rank of SR values for journal publications and conference proceedings for selected countries in 2007 16

Table 5: Document types in the SCIE and CPCI in the years 2000 - 2010 with at least 500 documents 17

Table 6: Relative annual growth in publication numbers for conference proceedings separately and in combination with journal articles 19

Table 7: Conferences that appear in the top 10 publication sources in the CPCI from 2000 to 2010 21

Table 8: Coverage of the top 10 conferences in WoS 23

Table 9: Top conferences in the fields Computer Science, Engineering, Medicine and Multidisciplinary from 2000 to 2010 25

Table 10: Coverage of the top conferences in field of Computer Science, Engineering, Medicine and Multidisciplinary in the CPCI from 2000 to 2010 27

Table 11: Share of conference proceedings in the overall publication output 28

Table 12: Growth of publications since 2000 measured for the year 2009 for conference proceedings and journal publications 30

Table 13: Percentage of citations received after 5 years that was already emitted in a 3 year citation window 31

Table 14: Citation rates for journal publications, conference proceedings and all documents in the year 2006 32

Table 15: Citation rates in 27 scientific fields for journal publications, conference proceedings and all documents in the year 2006 33

Table 16: Percentage of publications that have been republished as journal documents under the same title 35

Table 17: Percentage of conference proceedings in resubmissions 36

Table 18: Number of sources/publications in SCIE that are associated with a publication source title containing one of our keywords.....37

Table 19: Shares and ranks of selected countries with in- or exclusion of conference proceedings within all publications in 2009 38

Table 20: Ranks of selected countries according to citation rate with in- or exclusion of conference proceedings in 2006 39

Table 21: Citation rate of selected countries for different document type combinations in 2006 40

Table 22: Ranks of selected countries with in- or exclusion of conference proceedings in the fields Physics, Engineering and Computer Science in 2009.....43

Table 23: Ranks of selected countries with in- or exclusion of conference proceedings in the fields Electrical engineering, Measuring, Control and Optics in 2009.....46

Table 24: Changes in Germany's specialization with in- or exclusion of conference proceedings in 2009 47

Table 25: Changes in China's specialization with in- or exclusion of conference proceedings in 2009 48

Table 26: Changes in Japan's specialization with in- or exclusion of conference proceedings in 2009 50

Table A1: Changes in South Korea's specialization with in- or exclusion of conference proceedings in 2009 56

Table A2: Changes in Finland's specialization with in- or exclusion of conference proceedings in 2009 57

Figures

Figure 1: Publication share for journal publications for selected countries from 2000 to 2010 9

Figure 2: Publication share for conference proceedings for selected countries from 2000 to 2010 10

Figure 3: Citation rate for journal publications for selected countries from 2000 to 2009 11

Figure 4: Citation rate for conference proceedings for selected countries from 2000 to 2009 12

Figure 5: IA for journal publications for selected countries from 2000 to 2009 13

Figure 6: IA for conference proceedings for selected countries from 2000 to 2009 14

Figure 7: SR for journal publications for selected countries from 2000 to 2009 15

Figure 8: SR for conference proceedings for selected countries from 2000 to 2009 16

Figure 9: Number of journal publications and conference proceedings in comparison 18

Figure 10: Number of different conferences per year covered in the CPCI 19

Figure 11: Distribution of journal publications and conference proceedings in the 27 fields in 2009 20

Figure 12: Citation flows of conference proceedings and journal publications in 2000 and 2010 in percent 34

Figure 13: Number of publications in Physics with in- or exclusion of conference proceedings in 2009 41

Figure 14: Number of publications in Mechanical engineering with in- or exclusion of conference proceedings in 2009 42

Figure 15: Number of publications in Computer Science with in- or exclusion of conference proceedings in 2009 42

Figure 16: Number of publications in Electrical engineering with in- or exclusion of conference proceedings in 2009 44

Figure 17: Number of publications in Measuring, control with in- or exclusion of conference proceedings in 2009 44

Figure 18: Number of publications in Optics with in- or exclusion of conference proceedings in 2009 45

1 Introduction/Aim of this study

At the end of 2008, Thomson Reuters extended the coverage of their publication database Web of Science (WoS) by conference proceedings. Up to that point, the publication database (officially) covered only journal articles. The introduction of the so called "Conference Proceedings Citation Indexes for Science and for the Social Sciences and Humanities" (CPCI) in 2008 allowed the access to conference proceedings in the WoS systematically. Thus, the combined (or separate) usage of publications in journals as well as in proceedings was enabled for bibliometric analyses on all scales. As for all bibliometric methods and analyses, the implications of different publication and citation behaviour as well as database coverage were left to the end user. In particular, Moed and Visser already mentioned in advance, that such an extension should account in any form for the fact that conference proceedings might be republished with only minor changes in other publication outlets (Moed/Visser 2007).

Even though it is not necessarily the database provider's responsibility to provide hints and means to all possible pitfalls of their data – corresponding to the "Caution hot!" warnings on all coffee cups since 1992 – the ignorance on the end-users side should be abolished by sufficient information on the "correct" and appropriate usage of the data. This is especially important in this case since false claims and deductions might actually hurt third parties, as they are the subject of the bibliometric studies and assessments.

The findings in this report indicate that the conference proceedings provide a (necessary) mean to extend the analysis of publications in particular for fields in which the dissemination of scientific results happens oftentimes at conferences. It can be fatal for subjects of bibliometric studies if conference proceedings are excluded from the analysis, because they are a comparably if not even essential substitute or alternative to journal publications at least in some fields. Thus, the complementary analysis of this publication type should be a standard in bibliometric studies.

Nonetheless, even if such a corresponding study is conducted, the publication types in the database might impede an accurate analysis. In general, the assignment of document types is not always correct and might lead to distorted results.[1] Even though Thomson Reuters introduced the document type "proceedings", this was not used for publications in the time span 2000 – 2010. Conference proceedings are thus also labeled as "articles" – as are genuine journal articles. Therefore, the distinction between

1 "Working with ISI data: Beware of Categorisation Problems", Anne-Wil Harzing, http://www.harzing.com/ISI_categories.htm, last accessed on 2012/10/18

those two types can only be made by the database products. Thus end-users who combine various products in the WoS have to distinguish between conference proceedings and articles via the product code.

However, even with this intermediate step, some conference proceedings are included in the other WoS products and can therefore not be detected automatically but only by a search for certain terms in the title of the source. The accuracy of such a method is difficult to assess. Also, it is unclear why specifically these proceedings were included. It could be assumed that these particular proceedings might deviate from others, e.g. in their standing or impact in the scientific community so that a coverage "among" articles is rectified. Furthermore, one has to ask whether the cost-benefit-ratio allows an elaborated detection of these hidden proceedings if the error they introduce in the bibliometric analyses is negligible.

The last part of this report describes how conference proceedings in general influence various types of bibliometric indicators. We take a look at different countries and analyze whether they profit or suffer from the inclusion of conference proceedings. Also, we compare the indicator values for various fields. This results in an overview of types of analyses, for which an in- or exclusion of conference proceedings might be advisable or not.

The remainder of this report is structured as follows: Section 2 gives an overview of the state of the art and compares literature that has already dealt with the database coverage of conference proceedings and their influence on bibliometric indicators. Furthermore, the approach taken in this report and necessary definitions are explained. To better illustrate the different results of bibliometric indicators, we calculate different bibliometric indicators for five countries and compare the results for conference proceedings and journal publications in Section 3. Section 4 consists of two parts. The first part, "Descriptive Analysis", shows the results for the quantitative assessment of the proceedings in the database. Using these findings as a basis, we then deduce the "Implications for Bibliometric Analyses" in the second part. The final section summarizes our findings and shows possibilities for future work.

2 Methodology

2.1 State of the Art

Importance of conference proceedings

The role of publishing new findings and methodologies at conferences differs in the fields of science. Thus, the various fields are represented to different extents in the set of conference proceedings. Kademani et al. (2009) found that publications in Physics and Engineering account for 52% and 22% respectively of the conference proceedings in Scopus. However, his results were restricted to scientists from a specific research center. According to Lisée et al. (2008), the relevance of conference proceedings is diminishing in most disciplines but not so in Computer Science and Engineering. The results by Eckmann et al. corroborate this, since 34 out of 65 Computer Vision researchers thought that a foregoing conference publication improves the chances for a journal publication (Eckmann et al. 2012). In addition to that, Glänzel et al. (2006) found that about one third of geosciences and more than 20% of physics, agriculture and mathematics is covered in the online WoS (both SCIE and CPCI[2]) in the period 1994-2002. This indicates that the ISI Proceedings are a valuable supplement to the WoS. Cocosila (2011) proved a general increase in the number of papers and a steady increase in cooperation at three conferences in the management information systems (MIS) field. Furthermore the leading MIS conference contributors tend to establish loyalty to a limited number of academic meetings. This shows the supplementary role of conference publications in scientific communication as well.

Another mentioned advantage of a conference proceeding regardless of later publications was the shorter time span between submission and publication date (Eckmann et al. 2012). However, journal articles seem to have a higher longevity since they are cited on average after 14.2 years, while the average citation age of conference proceedings equals 10.3 years (Lisée et al. 2008).[3] Franceschet (2010) also noted that conferences provide fast and regular publication of papers and help to bring researchers together. However, the impact of journal articles is significantly higher than the impact of conference papers in computer science; a severe skewness exists in the citation distribution for computer science papers. Thus, the mean citedness does not ap-

---

2 See definitions in Section 2.2.

3 These statements could not be corroborated with our findings though. The average citation age of articles lay between 5.6 (in 2000) and 2.7 (in 2007) years, while that for conference proceedings diminished from 4.9 (in 2000) to 2.6 (in 2007) years.

propriately reflect the impact of a conference publication in contrast to journal publications. Summing them up or mixing them is therefore not appropriate.

The importance of conference proceedings can also be measured by the number of citations they receive or emit. It reflects their role in communication and their impact in the scientific community. Again, their importance and thus their citations might vary among the different scientific fields. In Computer Science, Lisée et al. (2008) found that 20% of all citations refer to conference proceedings. Kademani et al. (2009) report that 45% of the citations of conference proceedings refer to conference proceedings in Physics. On the other hand, Lv et al. (2011) retrieved publications related to graphene in the SCI and CPCI database, and found that conferences proceedings are more concentrated than journal publications on the application of graphene, and that most highly cited papers in the graphene field are all conference proceedings. On the other hand, 26% of all citations in the WoS in 1996-2007 stem from conference proceedings (Meho/Rogers 2008). Accordingly, Bar-Ilan (2010) notes that even though the most highly cited publications in Computer Science are published in journals and journal publications receive more citations on average, a vast amount of theses citations is emitted from conference proceedings.

Furthermore, the role of conferences and scientific meetings has been underlined as an important indicator of the dynamic development in the field of Spanish social psychology (Íñiguez-Rueda et al. 2008). Hofer et al (2010) also pointed out that as a new and innovative, not very common approach, scientometric investigations of conference proceedings may be interpreted as early indicators of scientific development based on three clusters: the core, the semi-periphery and the periphery in international business field.

Resubmissions

Regardless of the advantages and disadvantages of conference proceedings, there is another issue when using them in bibliometric analyses that has been already addressed in the literature; conference proceedings are sometimes only a pre-version of a journal article so that the same work is basically published twice. Even though the authors might alter the article to some extent before the resubmission, the bibliometric consequences of this are similar to fraudulent duplicate submissions or Salami publishing (cf. Abraham 2000). It has been shown that this accounts for at least 22% if not more of the journal articles (Bar-Ilan 2010; Eckmann et al. 2012), which were published as a conference paper prior to the journal article. Looking at the other direction, in an early study by Drott (1995), 5 out of 32 conference proceedings were followed by a journal submission. Montesi and Owen show that in 5 journals in Software Engineering,

the share of extended versions of former conference proceedings lies between 3.8 and 58.1% (Montesi/Owen 2008). Aleixandre-Benavent et al (2009) found that 31.7% of all the papers presented at the 5th, 8th and 10th Conferences of the International Society for Scientometrics and Informetrics (ISSI) were republished. The median time lag between the conference publication and the journal publication varies between 2 and 4 years.[4]

It can be argued that republications might push certain bibliometric indicators, like for example simply the number of publications. This can be the case if the citations are not spread among the publications of similar work but rather doubled to refer to both publications; as a matter of fact Eckmann et al. (2012) show that in their sample, journal publications with a preceding conference publication are cited even more often than stand-alone publications. This could be the case because of the findings were already disseminated via the conference or because the resubmission was motivated by a high impact of the conference proceedings publication (which was also higher than of those without follow-up publications). Eckmann et al. (2012) argue that the authors of the conference paper "realized that those papers were receiving many citations ... and decided on improving those results for a second publication", which then in turn even had a yet higher impact. Furthermore, the journal publications receive more citations than their conference "predecessor". It might thus be asked whether both variants of the duplicate accounting for basically the same work should be included in a bibliometric analysis or not.

The survey of Montesi and Owen among the authors of the resubmissions shows which parts have been changed for the journal publication (Montesi/Owen 2008). It can be deduced that at least the main body was considerably or completely changed in approximately 60% of the articles. Also, in 46% of the articles, the revision took account for the discussion session at the previous conference. These findings corroborate the impression especially in Computer Science, that conference proceedings are merely a preceding step for a journal submission so that the suitability of a proposed approach, application or method can be discussed with the scientific community. If that was truly the case, conference proceedings should be excluded from the bibliometric analysis.

4 A finding that we were able to corroborate in Section 4.1.

2.2 Proposed Approach

This study is restricted to the WoS as licensed by the Komepetenzzentrum Bibliometrie. Different products are implemented in this version of the WoS so that they can be combined. Nonetheless, the origin of each document in respect to its product is still traceable. In this report, we divide the database in two parts: 1) The products that mainly cover conference proceedings and 2) those that deal with the remaining bibliometric data. The first part covers the subproducts "Social Sciences and Humanities Proceedings" and the "Scientific and Technical Proceedings" and will be compositely addressed by the term *CPCI* in the following. The remainder, namely the products SSCI, AHCI and SCIE, will be denoted as *SCIE* for the sake of simplicity since this report only distinguishes between conference proceedings and journal articles.

The whole analysis is restricted to the publication period between 2000 and 2010. This allows a trend analysis while taking into account that previous or later years might not be sufficiently covered (yet).[5]

We define the set of "journal publications" as those publications in the SCIE that are not included in the CPCI and which have the document type article, letter, review and note. "Conference proceedings" on the other hand encompasses all publications in the CPCI of the type article.[6] As we explain in Section 4.1, the document type "conference proceedings" is not applicable for our purpose. In some occasions, we deviate from this definition, in particular when looking for "hidden" conference proceedings in the SCIE. We point out such search strategies at the appropriate part of the text. "Hidden" documents are found by a text based search on the title of the publication source. All sources in the SCIE that have a title containing the terms "proceedings" or "conference" or starting with "proc." are marked as proceedings. A manually assessment of a sample of these source titles was performed to verify the applicability of the approach. Nonetheless, because of the restrictive nature of this approach and name variations, the results can be merely seen as a lower bound for the "gray mass" of hidden proceedings.

When not stated otherwise, the citations are calculated on the basis of all documents in a fixed citation window. We use citation windows of 3 and 5 years to compare the cita-

---

5 However, we will see later that even with an update in Spring 2012, the year 2010 is not completely covered in the database.

6 Additionally, to parallel the search strategy for the journal publications, we included the document types review, letter and note in the CPCI as well. In practice, this results in 10 more documents (all reviews) in the whole observation period.

tion behaviour for the two document types. For this comparison, the latest year that can be used is the year 2008, since the 5-year citation window spans to the year 2012.

Furthermore, our methodology for detecting resubmissions of former publications should be explained. Like Bar-Ilan (2010) we use matches of the document titles to identify duplicate publications. Also, at least one of the authors of both papers should be the same and they should have different IDs in WoS and different publication source titles. The resubmission might be published in the same or a following year.[7] For the calculation of conference proceeding resubmission, the first publication has to fulfill our definition of "conference proceedings" as given above and the second publication must correspond to a "journal publication". The numbers provided in this report always refer to the number of papers that were later resubmitted. Of course, the numbers decrease over time, since the available time for a resubmission diminishes for later years. If there is a pair of publications sharing the same title, we count it only once. If there are more than two publications *n* sharing the same title, we count the number of resubmissions (in total *n-1*).

Finally, "International Alignment (IA)" and "Scientific Regard (SR)" are used to evaluate publication activities for countries. IA describes whether the authors of a country release their achievements in internationally more or less visible journals, compared to the world average. SR describes whether the average citation rate of a country is higher or lower than the journal-specific expected citation rate, resulting from the average citation rate of journals in which the country's authors published their papers. Positive values of IA and SR show above-average level; negative values show below-average level; values of 0 is regarded equivalent to the average.

The IA value is calculated as follows:

$$IA_k = 100 \tanh \ln (EXP_k/OBS_w)$$

Here OBS_w denotes the actual observed citation rate of all publications in the world. EXP_k denotes the expected citation rate of the journals where the authors of this country published their papers.

---

7 In the case of equal publication years it cannot be distinguished between original publication and resubmission. This does not affect our analysis overall, because basically we know that both submissions were published by the same author with the same title and are thus duplicates and that is all we need. Nonetheless, we had to take this into account when calculating the numbers of resubmitted publications because otherwise the two publications of the same document type would have been counted as resubmitted publications as well as resubmissions.

The SR value is calculated as follows:

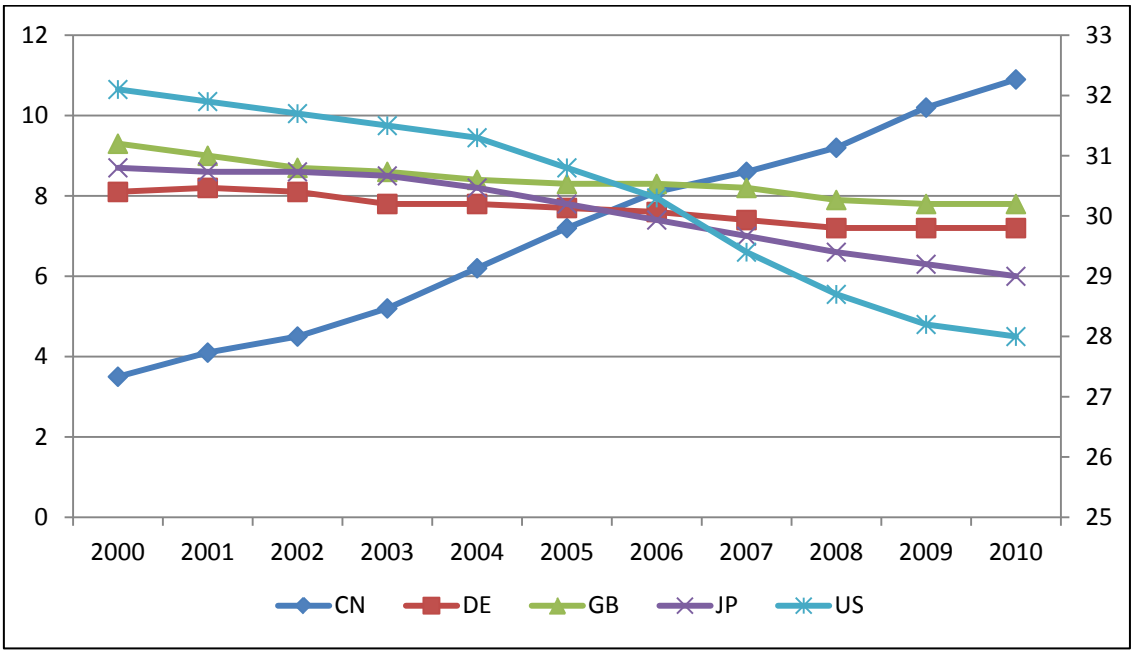
$$SR_k = 100 \tanh \ln (OBS_k/EXP_k)$$

Where OBS_k denotes the actual observed citation frequency of publications of country k. EXP_k denotes the expected citation rate resulting from the average citation frequency of the journals where the authors of this country published their papers.

3 Exemplary comparison of bibliometric indicators for selected countries

In this section, five selected countries, the US, Germany, Great Britain, Japan and China, are further investigated in order to analyze their trends of publication activities in terms of the two kinds of publication types.

Figure 1: Publication share for journal publications for selected countries from 2000 to 2010

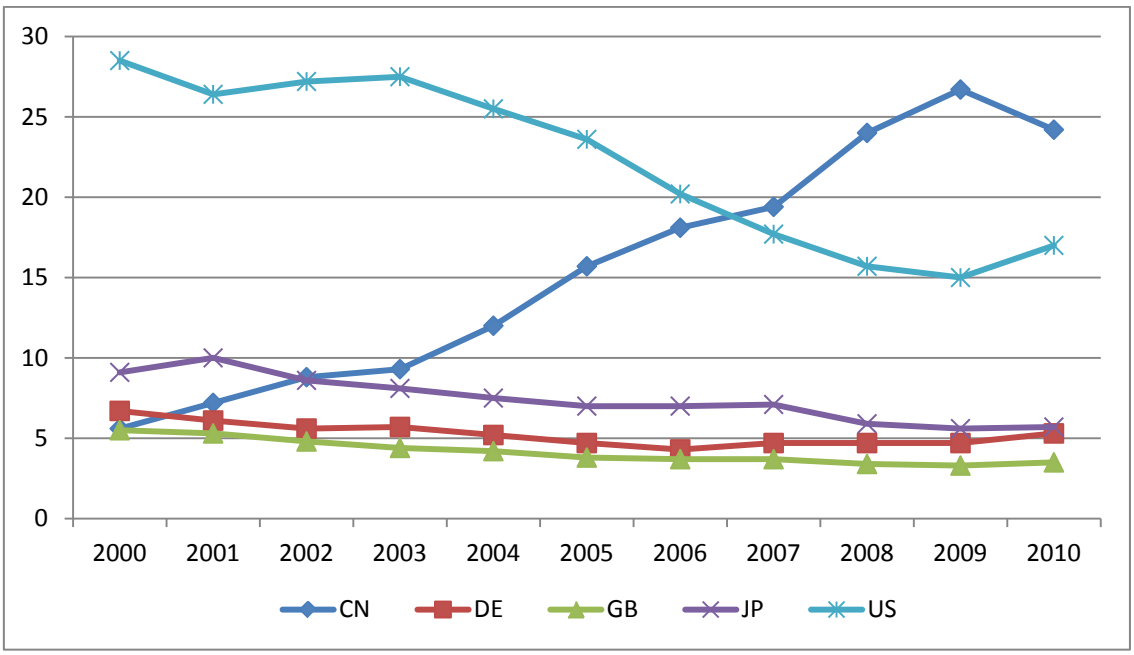


Source: SCIE; own calculations and illustration.

Figure 1 shows the shares of the five selected countries in worldwide publications for journal publications from 2000 to 2010, where it can be seen that the publication share of the US has decreased from 32.1% to 28.0% during the past decade. Similar decreases in percentages also can be found in Great Britain, Germany, and Japan; while for Japan the decrease is the highest. On the other hand, China's share has increased continuously and tremendously up to 10.9% in 2010.

Similar decreases and increases for the five selected countries can be observed in Figure 2, which displays the publication share for conference proceedings. China has an extraordinarily high growth rate for conference proceedings as well, but it is even higher than that for journal publications. Japan has a moderately decreasing share for conference proceedings compared with journal publications.

Figure 2: Publication share (%) for conference proceedings for selected countries from 2000 to 2010



Source: CPCI; own calculations and illustration.

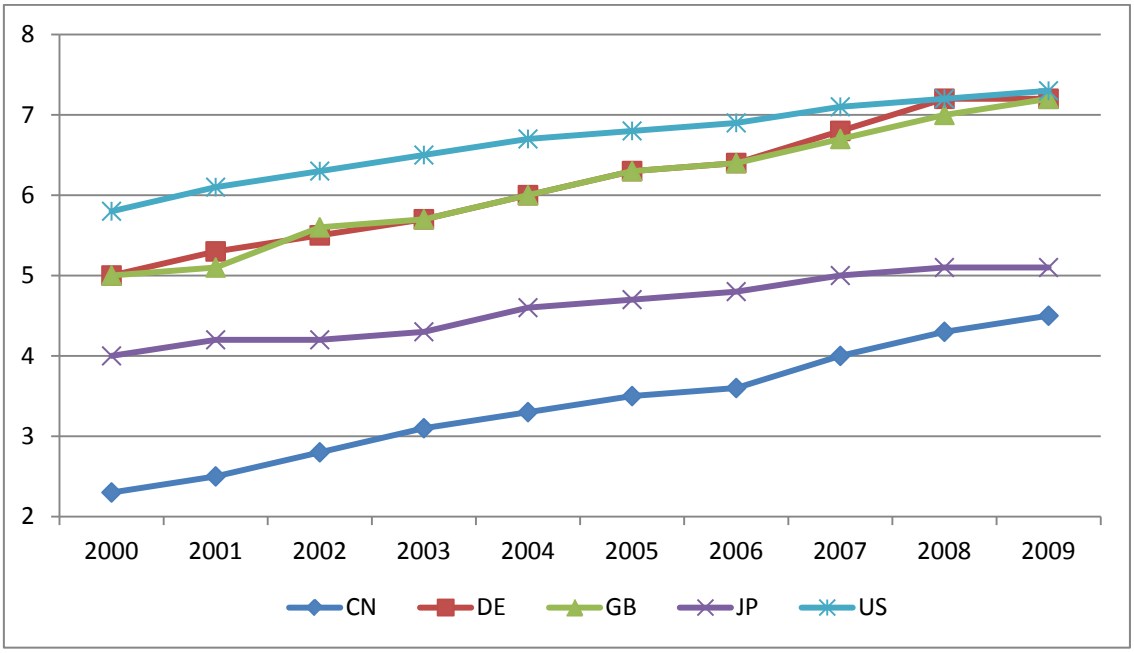
Table 1: Rank of publication shares (%) for journal publications and conference proceedings for selected countries in 2009

Country	Journal publications		Conference proceedings	
	Share	Rank	Share	Rank
CN	10.2	2	26.7	1
DE	7.2	4	4.7	4
GB	7.8	3	3.3	5
JP	6.3	5	5.6	3
US	28.2	1	15.0	2

Source: SCIE/CPCI; own calculations.

Ranks of publication shares for journal publications and conferences proceedings for the selected countries are shown in Table 1. It can be seen that the US rank first in terms of its publication share of journal papers, with a value of 28.2% in 2009; while China has exceeded the US and ranks first based on the share of proceedings (26.7%). In Japan, the absolute share of journal publications and conference proceedings is at a similar level, but the respective rankings differ.

Figure 3: Citation rate for journal publications for selected countries from 2000 to 2009

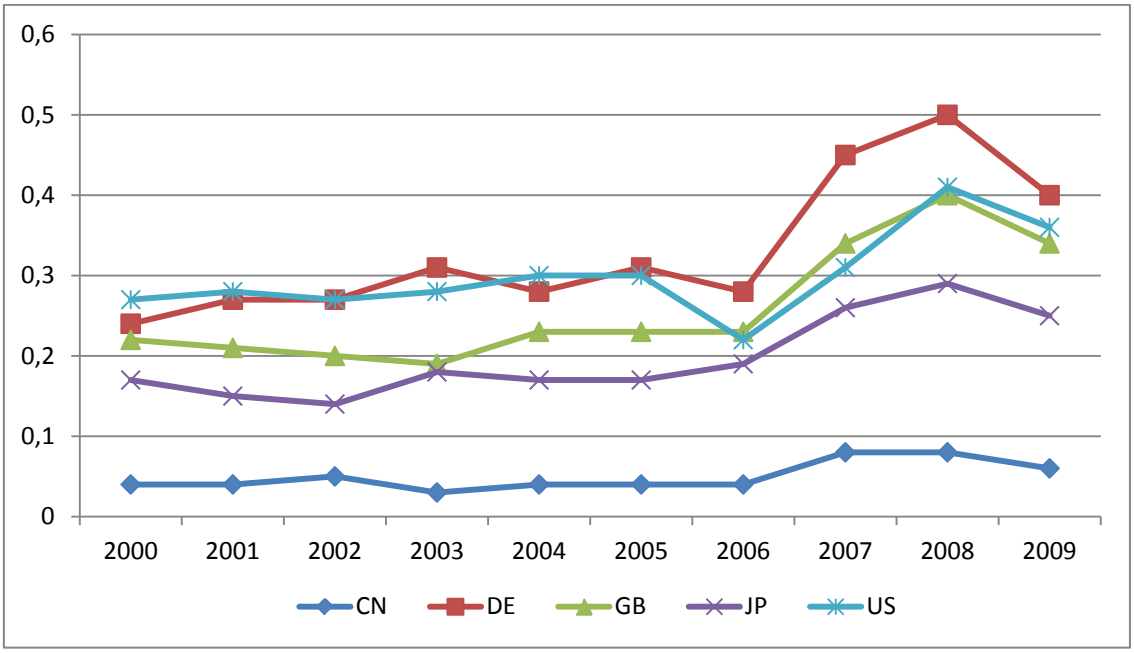


Source: SCIE; own calculations and illustration.

Trends of citation rates are observed further among the five selected countries. As shown in Figure 3, all selected countries maintain steadily upward trends in terms of the values of their citation rate for journal publications. At the same time, it is noted that the US, Germany and Great Britain reach a similar level, with a value of 7.2 in latest years. On the other hand, the Chinese citation rate for journal publications increased by a value of 2.2 during the past 10 years, but it still cannot catch up with those of industrialized countries.

As for the citation rate for conference proceedings, Figure 4 shows much smaller values compared with those for journal publications for all selected countries, with values of less than 0.5. Especially China's citation rates for proceedings alternate between 0.03-0.08 during the past ten years. All selected countries show a notable increase after 2006 in citation rates for conference proceedings, which may result from the introduction of the new database CPCI.

Figure 4: Citation rate for conference proceedings for selected countries from 2000 to 2009



Source: CPCI; own calculations and illustration.

Table 2: Rank of citation rate for journal publications and conference proceedings for selected countries in 2007

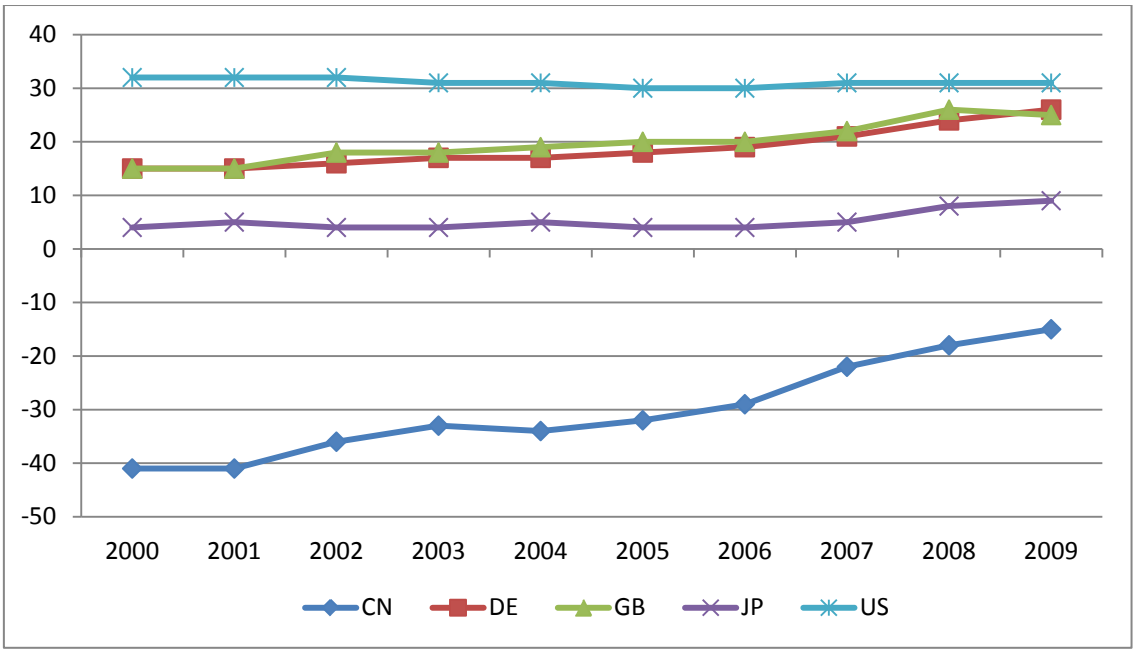
Country	Journal publications		Conference proceedings	
	Citation rate	Rank	Citation rate	Rank
CN	4.0	5	0.08	5
DE	6.8	2	0.45	1
GB	6.7	3	0.34	2
JP	5.0	4	0.26	4
US	7.1	1	0.31	3

Source: SCIE/CPCI; own calculations.

As to the comparison of citation rates for two kinds of publication types, it can be seen that the values for journal publications are much higher than for conference proceedings. It has to be remarked that the overall low citation rates for conference proceedings make a real comparison rather difficult. Nonetheless, Germany ranks first in terms of citation rate for conference proceedings, with a value of 0.45; while its rank for journal publications slips one spot to two, after the US, with a value of 7.1 (Table 2). Both citation rates in China are ranked last, and its value for conference proceedings is extraordinarily low.

Furthermore, IA values for journal publications are observed for five selected countries in Figure 5. The US have stable and top IA values in the past ten years among selected countries. While the IA-index for Germany as well as for Great Britain have improved tremendously from 15 to 25 in the last years, implying that both countries' scientists have increasingly preferred to provide their achievements for the international community in higher visible journals. China shows an increasing trend all the time and after 2007 even an obvious growth of its IA figures though it is still much lower than world average level, implying its efforts to enhance the scientific capacity and improve its international influence by rising shares of publications in higher impact journals.

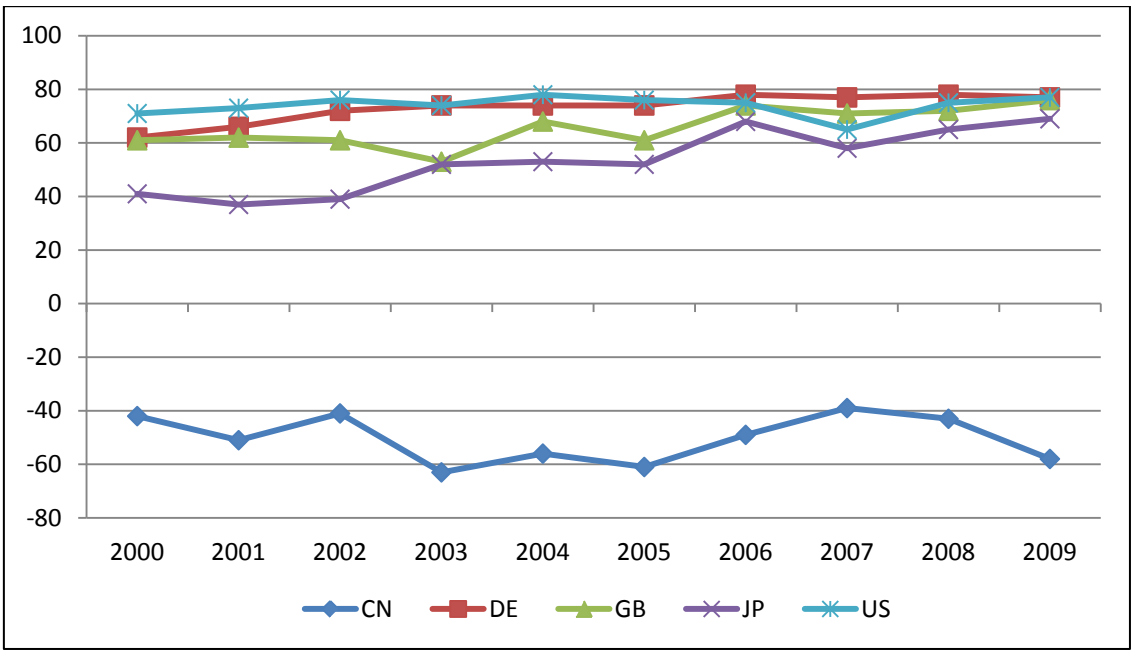
Figure 5: IA for journal publications for selected countries from 2000 to 2009



Source: SCIE; own calculations and illustration.

Again, IA value for conference proceedings are observed for the five selected countries. As shown in Figure 6, the US are not at a higher level than other countries for IA values of proceedings as it is for journal papers. Thus, Germany, Great Britain, the US and Japan have similar levels, with IA values of about 70 in 2009, implying above mentioned countries' scientists prefer to attend high international impact conferences, and publish their achievements there. However, it also can be seen that China has rather low IA values for proceedings, showing the smaller scientific impact of conferences that Chinese scientists have attended.

Figure 6: IA for conference proceedings for selected countries from 2000 to 2009



Source: CPCI; own calculations and illustration.

Table 3: Rank of IA values for journal publications and conference proceedings for selected countries in 2007

Country	Journal publications		Conference proceedings	
	IA	Rank	IA	Rank
CN	-22	5	-39	5
DE	21	3	77	1
GB	22	2	71	2
JP	5	4	58	4
US	31	1	65	3

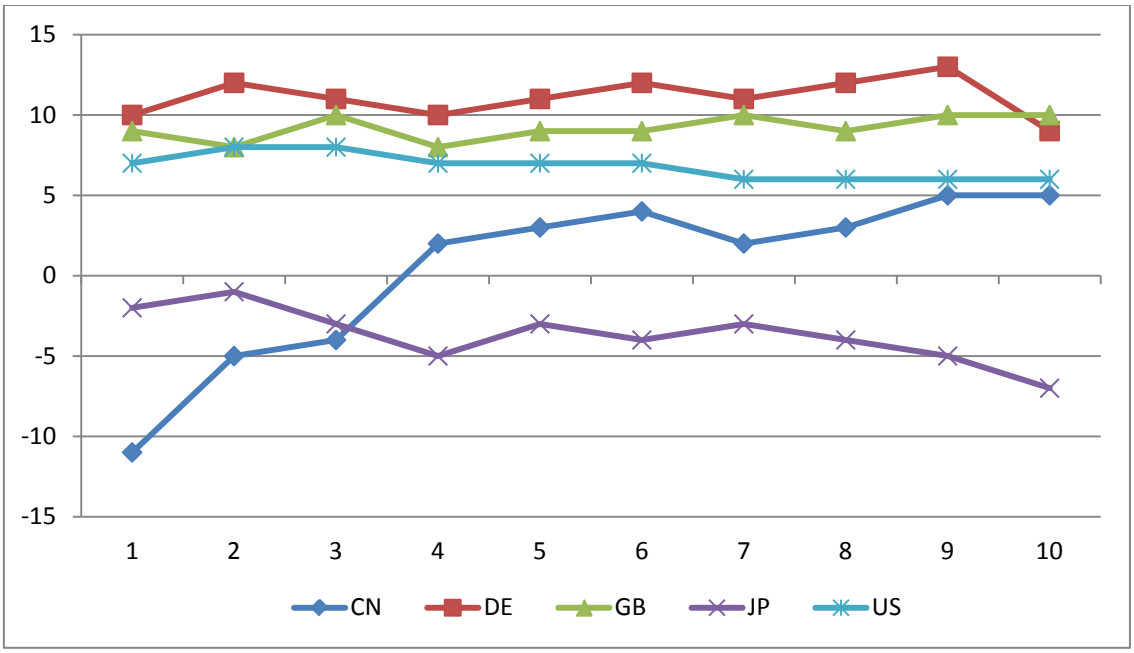
Source: SCIE/CPCI; own calculations.

Table 3 displays the comparison of IA values for two kinds of publication types. It can be seen that the US rank first before Germany and Great Britain in terms of IA for journal publications, which is contrary to the situation for conference proceedings, implying the journals in which the US scientists publish their papers have higher international impact; but the conferences that German and British scientists choose have even higher international impact. Ranks according to IA values for the two kinds of publication types are the same for China and Japan.

The scientific performances in selected countries are further inspected according to the journal-specific Scientific Regard (SR) for journal publications. As presented in Figure 7,

China shows a continuously increasing trend and rather high SR values in latest years, to a similar level like the US. However, since the IA index for China is much lower, it can be inferred from the result that China's publications were attracting citations at or above the expected values for their publications in lower impact journals. In comparison, Japan shows a continuous decrease in SR value in the past decade, which means that Japanese authors have published their papers in rather high impact journals, but received less citations than other papers in the same journals. Germany and Great Britain display a slight increase in SR value in the past ten years, while the US are on the opposite track, i.e. rather show a decreasing trend.

Figure 7: SR for journal publications for selected countries from 2000 to 2009

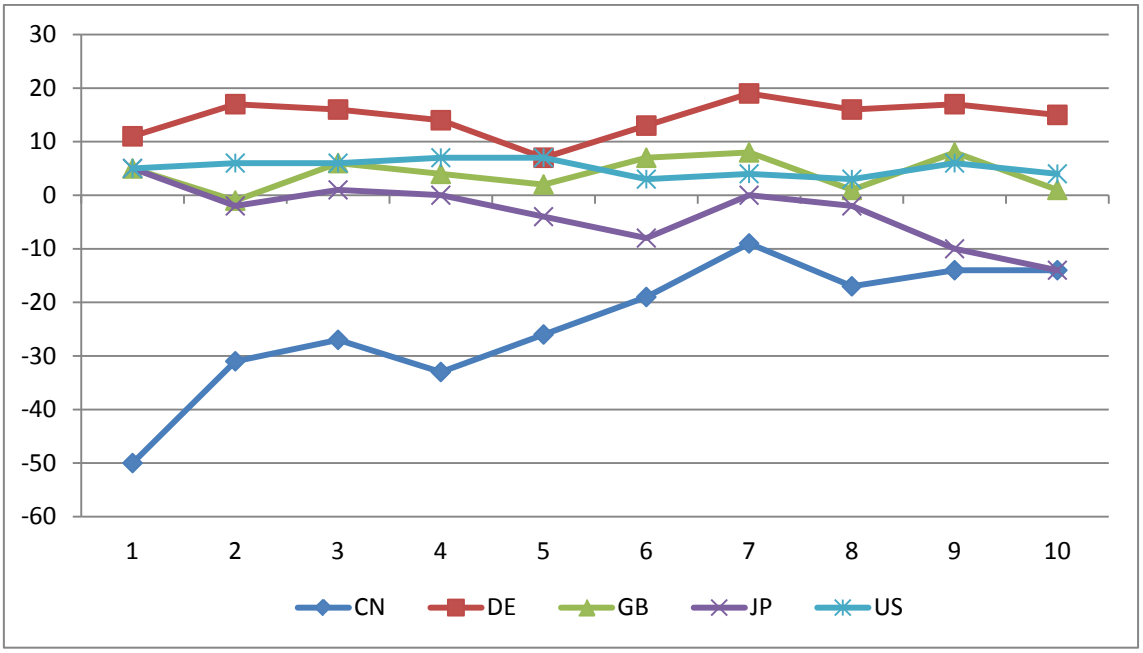


Source: SCIE; own calculations and illustration.

As to the SR values for conference proceedings for selected countries (Figure 8), Germany has always remained at a high level among the selected countries, which means German scientists have attended rather high impact conferences and attracted more attention than other papers in the same conferences. Though China's SR value for proceedings in the past years also grows remarkably, it is still much lower than world average level. It means that Chinese scientists have attended low scientific impact conferences, but their publications still received less attention than other papers at the same conferences. Similar to SR for journal publications, Japan also shows a continuous decreasing trend for proceedings, thus in both publication forms, Japan receives comparable less and less citations i.e. attention. This might be a result of higher inter-

national oriented conferences and journals that Japan chooses. So Japan aims higher in terms of quality but cannot (yet) keep up with its competitors.

Figure 8: SR for conference proceedings for selected countries from 2000 to 2009



Source: CPCI; own calculations and illustration.

Table 4: Rank of SR values for journal publications and conference proceedings for selected countries in 2007

Country	Journal publications		Conference proceedings	
	SR	Rank	SR	Rank
CN	3	4	-17	5
DE	12	1	16	1
GB	9	2	1	3
JP	-4	5	-2	4
US	6	3	3	2

Source: SCIE/CPCI; own calculations.

Table 4 shows the comparison of SR values between journal publication and conference proceedings. German SR figures for the two kinds of publication types have always ranked first in the past ten years, namely that German publications attract more citations expected for the journals and conferences in which they appear. The performance of SR index for journal publication for Great Britain is better than the US, while its performance for proceedings is comparatively worse. Japan and China similarly swap ranks when the publication type is changed.

4 Results

4.1 Descriptive Analysis

In this section, we will take a look at the quantitative assessment of the conference proceedings. That is, we will compare their share in the database with those of the journal publications. This provides a better understanding of the consequences of deciding for or against the inclusion of conference proceedings in a bibliometric analysis. Furthermore, the stability of coverage of the conference proceedings is discussed in terms of overall number and included sources. Finally, we compare the absolute and relative number of conference proceedings republished as journal articles with those in general.

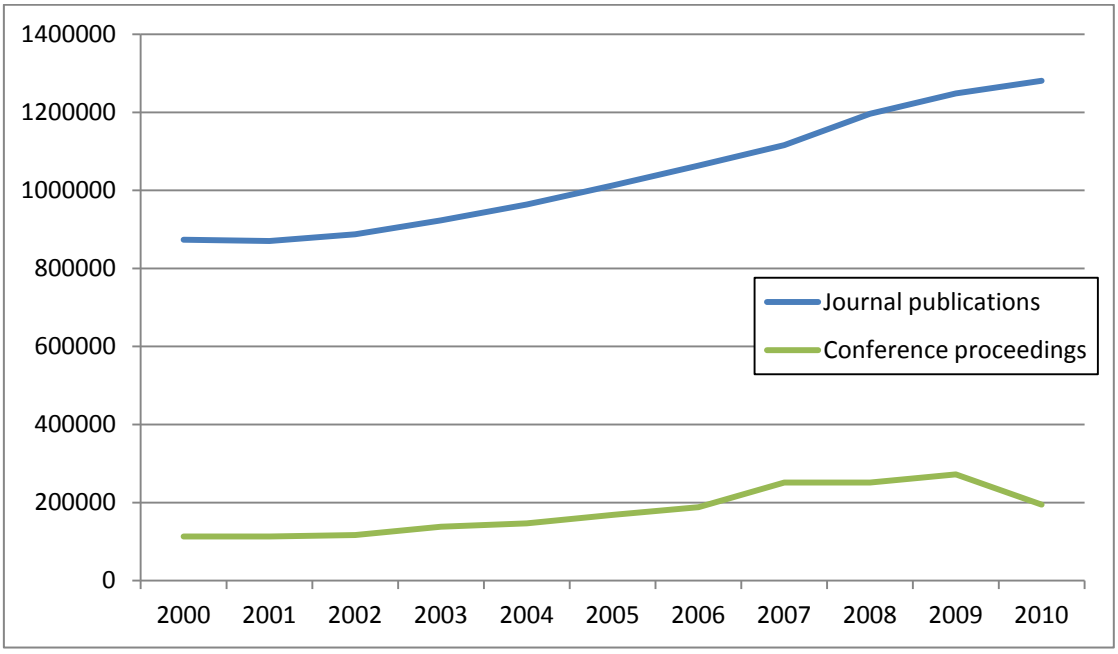
Table 5: Document types in the SCIE and CPCI in the years 2000 – 2010 with at least 500 documents

Document Type	SCIE	CPCI	Ratio (rounded)
News Item	274,502	0	--
Reprint	8,340	15	556:1
Bibliography	4,620	1	4,620:1
Article	10,445,518	1,954,323	5:1
Art Exhibit Review	30,505	0	--
Book Review	810,406	1	810,406:1
Correction, Addition	102,048	8	12,756:1
Editorial Material	780,817	206	3,790:1
Film Review	20,085	0	--
Music Score Review	4,227	0	--
Record Review	20,044	0	--
Biographical-Item	67,116	11	6,101:1
Music Performance Review	13,797	0	--
Music Performance Review	457,071	0	--
Meeting Abstract	2,061,494	7,671	269:1
Fiction, Creative Prose	9,963	0	--
Review	533,422	10	53,342:1
Theater Review	6,230	0	--
TV, Radio or Videocassette Review	4,686	0	--
Excerpt	1,578	0	--
Poetry	59,997	0	--
Dance Performance Review	5,754	0	--

Source: SCIE/CPCI; own calculations.

Firstly, Table 5 shows the different kind of document types used in the WoS. Document types that covered less than 500 documents in both database products in the time period were excluded. As already mentioned in Section 2.2, we use the document types article, review, letter and note only in this analysis. Other document types seem less important for the scientific research. The document type "proceeding paper" is not listed since it was not used in the WoS in the time period 2000 – 2010.[8] The majority of publications in the CPCI are assigned to the document type article. At the same time, it is noteworthy that proceedings recorded in the SCIE database reflect only conference proceedings published in journals (Sigogneau 2000), and Harzing (2013) noted that Thomson Reuters misclassified journal articles containing original research into the "review" or "proceedings paper" category. Above mentioned uncertainty may make research results by different groups less comparable.

Figure 9: Number of journal publications and conference proceedings in comparison



Source: SCIE/CPCI; own calculations and illustration.

Figure 9 compares the number of journal publications and conference proceedings in WoS. The ratio of conference publications and journal articles in this time span ranges from approximately 1:10 to 1:4 (not shown). The relative annual growth rate of the conference proceedings is much higher than those for the journal articles. Between 2006 and 2007, the conference publication numbers enhanced by 34%, but for 2008 there is

8 Note that we are referring to the version of the WoS that is sold for in-house usage by Thomson Reuters as stated in Section 2.2.

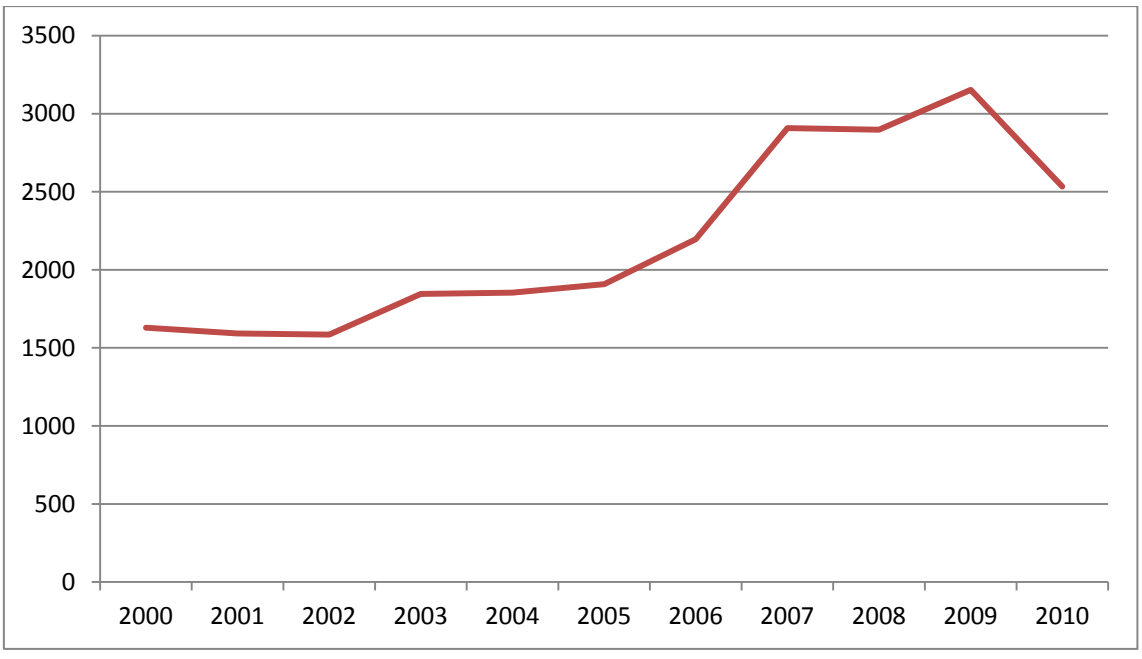
no notable increase in publication numbers (see Table 6).[9] This observation can probably best be explained by the introduction of the CPCI in 2008, for which a high coverage of the most recent publications seemed most important. For the other years between 2002 and 2009, we have varying growth rates between 3% (2002) and 18% (2003) for conference proceedings. For journal publications, the growth rate is rather stable around 6% for the same time span. The conference proceeding numbers collapse remarkably in the year 2010. This cannot be due to the fact that we are too close to the current edge because the journal publications seem to be sufficiently covered for this year. For the publication sources (in this case the conferences, see Figure 10) as well a substantial drop for the last year becomes visible. Apparently, fewer conferences for this year were included for unknown reasons.

Table 6: Relative annual growth in publication numbers for conference proceedings separately and in combination with journal articles (all)

Document type	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
All	0%	2%	6%	5%	6%	6%	9%	6%	5%	-3%
Conference proceedings	0%	3%	18%	6%	15%	12%	34%	0%	8%	-29%

Source: SCIE/CPCI; own calculations.

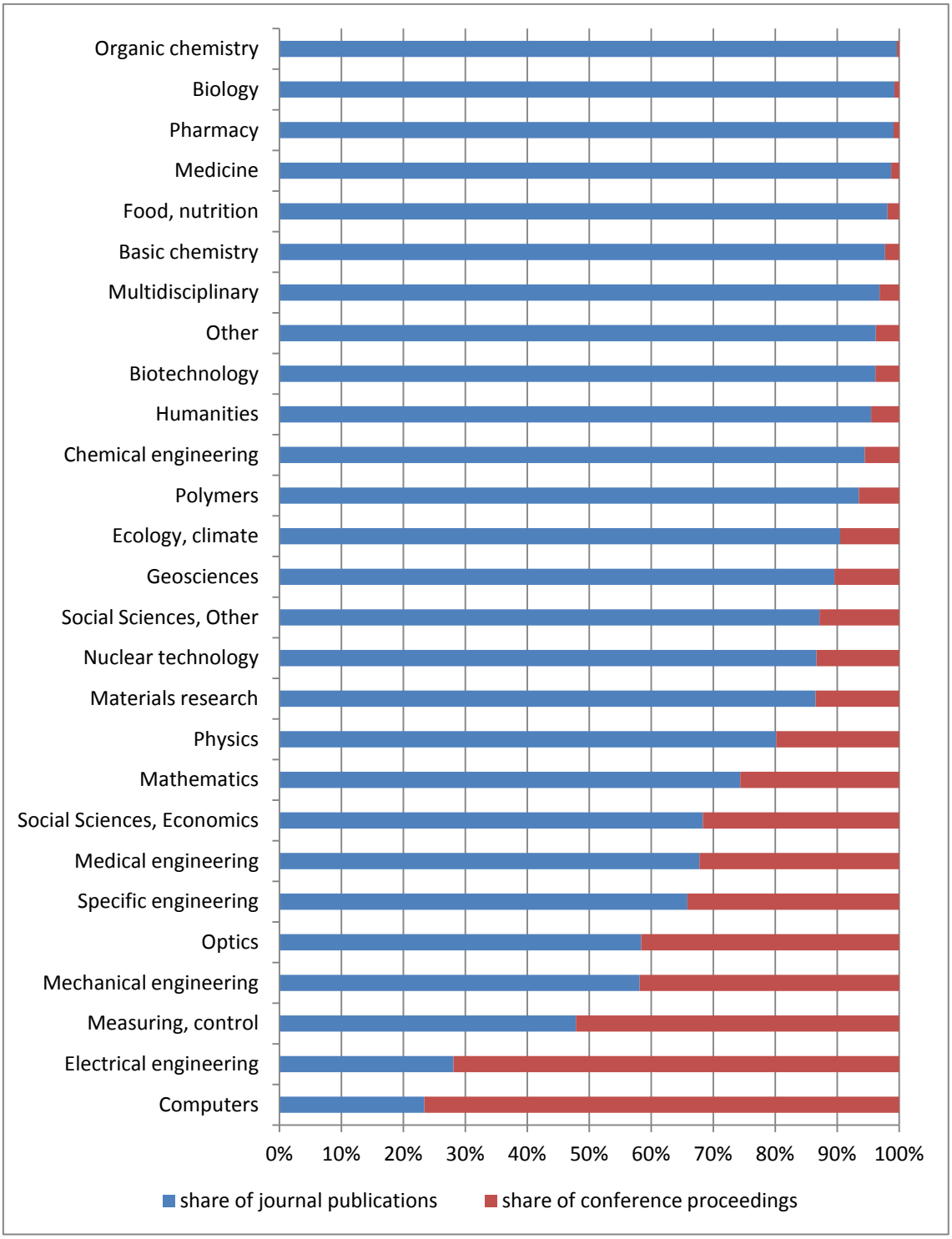
Figure 10: Number of different conferences per year covered in the CPCI



Source: CPCI; own calculations and illustration.

9 In absolute numbers, 2008 has 7 publications less than 2007.

Figure 11: Distribution of journal publications and conference proceedings in the 27 fields in 2009



Source: SCIE/CPCI; own calculations and illustration.

Table 7: Conferences that appear in the top 10 publication sources in the CPCI from 2000 to 2010

No.	Conference Name	No.	Conference Name
1	American Control Conference	19	IEEE Particle Accelerator Conference
2	ANTEC: Society Of Plastics Engineers	20	IEEE Power And Energy Society General Meeting (PESGM)
3	Asia Pacific Microwave Conference	21	IEEE Power Engineering Society Winter Meeting
4	Chinese Control And Decision Conference (CCDC)	22	IEEE VTS Vehicular Technology Conference
5	Conference On Lasers & Electro-Optics/Quantum Electronics And Laser Science Conference (CLEO/QELS)	23	IEEE/RSJ International Conference On Intelligent Robots And Systems
6	EBM: International Conference On Engineering And Business Management	24	IFMBE
7	ED-MEDIA: World Conference On Educational Multimedia, Hypermedia & Telecommunications	25	International Conference On Acoustics Speech And Signal Processing (ICASSP)
8	IEEE Antennas And Propagation Society International Symposium	26	International Conference On Bioinformatics And Biomedical Engineering
9	IEEE Conference On Decision And Control (CDC)	27	International Conference On Machine Learning And Cybernetics
10	IEEE Engineering In Medicine And Biology Society Conference	28	International Conference On Wireless Communications Networking And Mobile Computing (WICOM)
11	IEEE Global Telecommunications Conference (Globecom)	29	International Congress On Image And Signal Processing
12	IEEE International Conference On Communications	30	International Symposium On Test And Measurement (ISTM)
13	IEEE International Conference On Image Processing (ICIP)	31	Particle Accelerator Conference
14	IEEE International Conference On Robotics And Automation (ICRA)	32	Pacific Rim International Conference On Advanced Materials And Processing
15	IEEE International Conference On Systems Man And Cybernetics	33	SICE-ICASE International Joint Conference
16	IEEE International Symposium On Circuits And Systems	34	Springer Proceedings In Physics
17	IEEE International Symposium On Geoscience And Remote Sensing (IGARSS)	35	World Conference On Photovoltaic Energy Conversion, VOLSA-C
18	IEEE Nuclear Science Symposium Conference	36	World Congress On Intelligent Control And Automation (WCICA)

Source: CPCI; own calculations.

These numbers already hint at the high volatility of the conference proceedings coverage. However, to get the whole picture, we also refer to the numbers for the different

fields in science and also the coverage of individual conferences. Since the coverage in the year 2010 seems to be insufficient, we restrain to comparisons for the year 2009 if a single year analysis is necessary.

First, we compare the shares of conference proceedings in the 27 fields (Figure 11). The ratio of conference proceedings to journal publications shows the importance of the respective document type for the scientific communication in the fields. More than half of the publications in Computer Science, Electrical Engineering and Measuring, control are conference proceedings. Other more technical or engineering focused fields also have conference proceedings shares of more than 20%. These results already suggest that the impact of deciding for or against conference proceedings in an analysis varies among the fields.

Table 7 gives an overview of the set of the top 10 conferences in the years 2000 – 2010 in the WoS. Since there are some conferences appearing multiple times in the top 10 of the single years, we can reduce the set of investigation to 36 conferences. Still, the volatility in coverage and thus in the top 10 conferences reflects in the high number of different conferences appearing in this set.

The conferences were ordered alphabetically in Table 7 . The number of publications in the top 10 per year varied between 531 and 3061 per conference and year. The maximum value seems a rather high number of publications for one single conference, but indeed the proceedings of the conference No. 10 for example encompassed 7276 pages in the year 2009.

The list was manually revised since spelling variants were used in WoS. For instance, the conference No. 10 had the following label variations:

- PROCEEDINGS OF ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY
- 2006 28TH ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY, VOLS 1-15
- IEEE Engineering in Medicine and Biology Society Conference Proceedings
- EMBC: 2009 ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY, VOLS 1-20

Table 8: Coverage of the top 10 conferences in WoS (black=top 10 conference in that year, grey=covered in CPCI but not in top 10 list, white=not covered)

Source: CPCI; own calculations and illustration.

This short list also exemplifies the necessity for a case independent search strategy like the one we used here. A case independent search matches texts even if the case of the letters differs. In contrast, a case sensitive search with capital letters only would not return the third entry, neither would a case sensitive search with lowercase letters only. The latter implementation would even result in an empty result set. With a case independent search strategy we ascertain the maximum number of possible matches given a conference name. A matching by source ID was not possible, because a different source ID was assigned to each conference in each year. This aspect hinders the tracking of specific conferences even further.

Table 8 shows the coverage of the conferences of Table 7 over the years. A black cell marks a ranking among the top 10 conference in the WoS in that particular year in regard to publication numbers. Cells in grey refer to conferences that could be found by a query with the conference title in WoS, but were not listed in the top 10. For instance, conference No. 36 was only included every other year, but was not even covered in the year 2004 in WoS. In fact, in this case we are actually dealing with a biannual conference.[10] Since we had to query for these conferences manually, we can not be absolutely sure that they are not covered under some other name in the database.

10 see http://wcica12.amss.ac.cn/history, accessed on 2012/11/21

Table 9: Top conferences in the fields Computer Science, Engineering, Medicine and Multidisciplinary from 2000 to 2010

Field	No.	Conference Name	Publications according to Microsoft Academic Search	Publications in WoS
Computer Science	1	CVPR - Computer Vision and Pattern Recognition	7,731	3,107
	n/a	CHI - Computer Human Interaction	8,341	0
	2	INFOCOM - IEEE INFOCOM	6,557	3,889
	n/a	WWW - World Wide Web Conference Series	2,924	56(in 1999)
	3	NIPS - Neural Information Processing Systems	4,276	1,305
	4	KDD - Knowledge Discovery and Data Mining	2,061	251
	n/a	SIGMOD - International Conference on Management of Data	2,125	0
	n/a	VLDB - Very Large Data Bases	2,720	72(in 1996)
	5	MOBICOM - Mobile Computing and Networking	938	101
	6	SIGCOMM - ACM SIGCOMM Conference[11]	944	102
Engineering	7	VTC - IEEE Vehicular Technology Conference	19,394	4,876
	8	GLOBECOM - IEEE Global Telecommunications Conference	15,637	10,369
	9	ACC - American Control Conference	12,767	13,184
	10	IEEE MTT-S International Microwave Symposium	13,573	7,210
	11	IEEE Military Communications Conference	8,399	3,518
	12	APEC – Annual IEEE Conference on Applied Power Electronics Conference and Exposition	5,044	3,503
	13	ECTC - Electronic Components and Technology Conference	5,601	4,226
	n/a	ISSCC - IEEE International Solid-State Circuits Conference	7,487	0
	14	IEDM - International Electron Devices Meeting	1,0261	3,197
	15	IWUWBS - IEEE Conference on Ultra Wideband Systems and Technologies	255	102
Medicine	16	IEEE International Symposium on Biomedical Imaging	3,479	2,489
	n/a	MICCAI - Medical Image Computing and Computer-Assisted Intervention	2,955	0
Multi-disciplinary	17	ESM - European Simulation Multiconference	505	160

Source: CPCI, Microsoft Academic Search; own calculations.

11 A former workshop which evolved to a conference in 2002

A similar analysis was performed to find the conferences that are deemed important in the scientific community. In this case, we used an alternative source, external to WoS: Microsoft Academic Search[12] provided the top 10 conferences for the fields. Not all fields were covered or filled with up to 10 conferences. In Table 9 the conferences that were available are listed.[13] We can confirm the list of conferences for Medicine and Computer Science.

Those conferences in the list for which no equivalent could be found in WoS are marked with a "n/a" in the second column. A comparison of publication numbers according to both sources was intended, but unfortunately Microsoft only provides numbers for the past 10 or 5 years or in total. As was already mentioned, the insufficient coverage in the most recent years in WoS makes a comparison based on these very difficult. So we decided to use the total numbers in both databases. There is only one conference for which WoS has more entries than Microsoft (namely No.9). All other conferences have a lower number of entries. Six conferences are not covered at all in the WoS leading to a huge number of missing documents.

Similar to the previous analysis we found some discrepancies between the online version of the WoS and the version that is sold for in-house databases. For instance, a query for conference No. 2 resulted in 2734 entries between 1994 and 2010 in the online database[14]. 31 different kinds of spelling were used to denote the conference in that time period. In the in-house database, a query for all conference names containing the term "INFOCOM" at any position in the string and a manual selection afterwards produced four different name variants: "IEEE INFOCOM", "IEEE INFOCOM SERIES", "IEEE INFOCOM ’98 - THE CONFERENCE ON COMPUTER COMMUNICATIONS, VOLS. 1-3" and "IEEE INFOCOM 2009 - IEEE CONFERENCE ON COMPUTER COMMUNICATIONS WORKSHOPS". While the former was the most common (covering 3541 of 3999 publications), the latter two are good examples of the name variants used also in the online version. In the WoS online database, the conference name (source name) contained in most cases also the year or the sequence number of the conference in the respective year. Nonetheless, ambigue conference names like "IEEE INFOCOM CONFERENCE", "CONFERENCE ON IEEE INFOCOM" and "IEEE INFOCOM MEETING" can be found here as well. In the in-house database, the aforementioned spelling variant "IEEE INFOCOM 2009 [...]", which covers 110 publications,

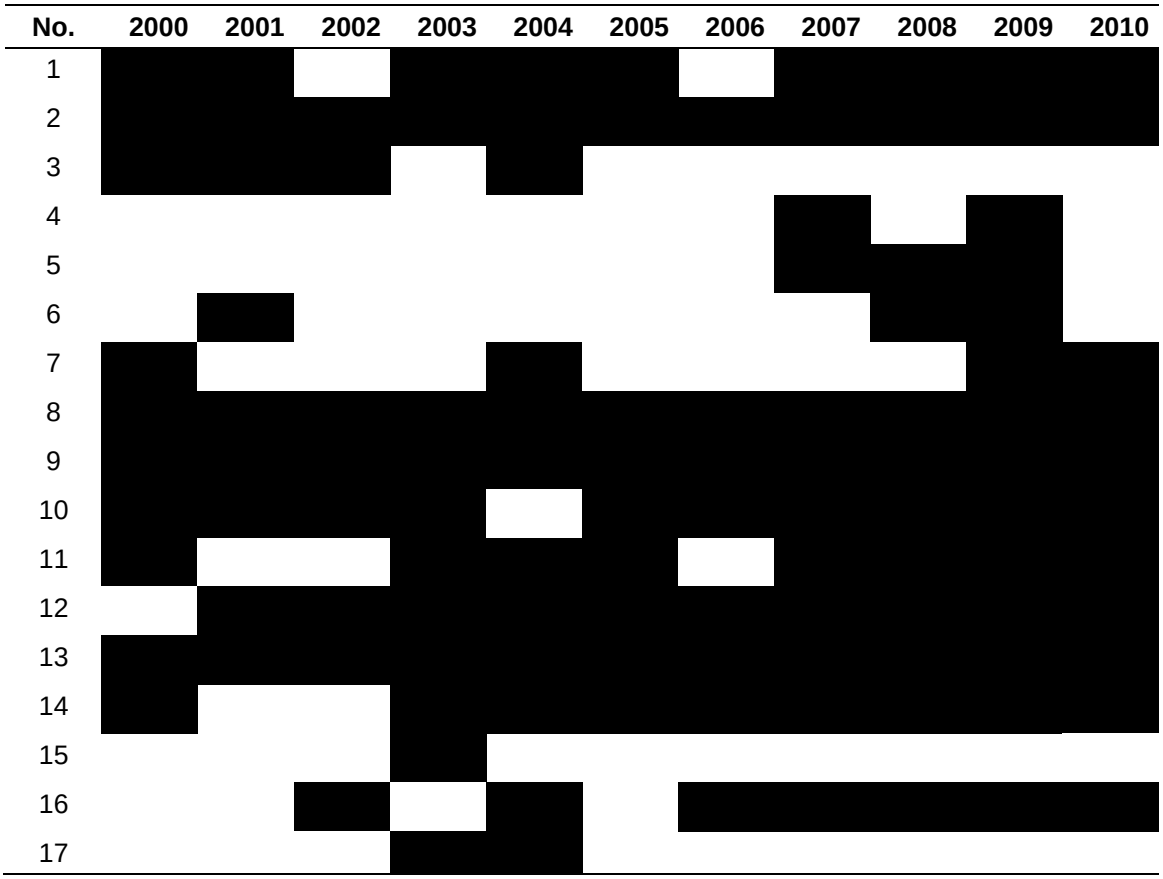
12 http://academic.research.microsoft.com/

13 website accessed 2012/11/27

14 In both the online database as in the in-house database, the search was performed on CPCI-SSH and CPCI-S.

is paralleled by 378 entries that are assigned to the arbitrary name "IEEE INFOCOM" but have the publication year 2009 as well. Thus it is unclear if some publications are included twice under different conference names or if these are distinct publication sets.

Table 10: Coverage of the top conferences in field of Computer Science, Engineering, Medicine and Multidisciplinary in the CPCI from 2000 to 2010 (black=covered in that year, white=not covered)



Source: CPCI; own calculations and illustration.

To complete the picture of the coverage, we also investigated for each conference the years of coverage in WoS. The search was performed analogously to the previous search for the top conferences in WoS. Again, we find major gaps in single years or conferences that are only covered temporarily (Table 10).

Thus, Table 8 and Table 10 show that even though there are some conferences that are covered in the whole time span other conferences make only sporadic appearances in the database for various reasons. As was already mentioned in the above example, some conferences have a biannual cycle. Yet other conferences are simply not covered, resulting in huge gaps in the coverage. Note the impact of the shortfall of

just one of these conferences which results in the loss of at least a few hundreds of publications for a field. Imagine the impact on individuals, who submitted to a conference that was in one year one of the top 10 conferences and in the next not even covered in the database. Similar impact can be perceived on the coverage of whole fields.

These observations make the already noted high volatility in both database forms even clearer and show the need for extensive semi-automatic queries in the case of single source analysis. A similar search for the journal "Scientometrics" in the time period 1980 to 2011 showed that in this case no differences between online and in-house database were detectable; for each year, the number of publications was exactly equal. Thus, this seems to be an issue that is connected with the publication type as well.

A very interesting part is conference No. 6 which started as a workshop in 2000 but reached the status of a conference by 2002. Interestingly, the second instance of the workshop in 2001 was covered in WoS but then again only the years 2008 and 2009 in addition, which in turn results in approximately 800 not covered documents.

Table 11: Share of conference proceedings in the overall publication output (journal publications and conference proceedings)

	2000	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	***2010***
BR	10%	11%	9%	10%	10%	9%	9%	9%	9%	9%	*6%*
CA	9%	8%	9%	10%	10%	10%	12%	13%	11%	10%	*8%*
CH	7%	8%	7%	8%	8%	8%	7%	10%	8%	8%	*7%*
CN	17%	19%	20%	21%	23%	27%	28%	34%	35%	36%	*25%*
DE	10%	9%	8%	10%	9%	9%	9%	13%	12%	13%	*10%*
DK	6%	5%	6%	6%	6%	7%	6%	9%	7%	7%	*5%*
ES	7%	8%	7%	8%	9%	9%	9%	13%	11%	12%	*9%*
FI	11%	10%	11%	11%	12%	12%	12%	15%	13%	14%	*9%*
FR	8%	8%	8%	9%	9%	9%	10%	13%	12%	12%	*9%*
GB	7%	7%	7%	7%	7%	7%	7%	9%	8%	8%	*6%*
IN	6%	4%	6%	6%	7%	7%	11%	10%	11%	12%	*8%*
IT	11%	11%	11%	12%	11%	12%	12%	14%	12%	13%	*10%*
JP	12%	13%	12%	13%	12%	13%	14%	19%	16%	16%	*13%*
KR	14%	15%	11%	13%	12%	14%	18%	27%	18%	14%	*9%*
NL	9%	8%	8%	8%	8%	8%	8%	10%	9%	9%	*6%*
RU	9%	8%	9%	9%	10%	10%	10%	10%	8%	8%	*6%*
SE	8%	7%	6%	7%	7%	8%	8%	9%	8%	8%	*6%*
US	10%	10%	10%	12%	11%	11%	11%	12%	10%	10%	*8%*
ZA	8%	7%	8%	6%	7%	6%	7%	9%	7%	8%	*5%*
WORLD	11%	12%	12%	13%	13%	14%	15%	18%	17%	18%	*13%*

Source: SCIE/CPCI; own calculations.

Next, we take a look at the role of proceedings for different countries. Table 11 shows the share of conference proceedings in each country's overall publication output. Again, we see a massive drop in conference proceedings numbers and thus the shares for 2010 (marked in italics). Even though the total share of conference proceedings amounts for 18% in 2009, the individual shares for the selected countries vary between 7 and 16% - except for China which shows a tremendous share of conference proceedings of about 36%. China, which is among the top publishing nations in our set (cf. Section 4.2), majorly influences the overall proceeding share that was already depicted in Figure 9. This huge share of conference proceedings in 2009 is on the one hand a product of an already huge share in 2000 which was on the other hand increased tremendously during the time period 2000 – 2009. In the end, nearly one third of all publications in China are conference proceedings. One possible explanation could be the increasing number of conferences hosted in China. Some top conferences now have "little brothers" in Asia, for which the number of Asian attendances might be higher than in other countries (as is the case for all conferences concerning the relatedness of proximity to and visitors from other countries). For instance, the ISWC (International Semantic Web Conference) ranks as one of the top conferences in Computer Science. Held for the first time in 2002, a European counterpart[15] was initiated in 2004 that was followed by an Asian version in 2006 (ASWC, Asian Semantic Web Conference, Beijing, China). Similar observations can be made for other conferences as well. The localization of such conferences might especially encourage younger researchers to participate because of lower travel costs or less time demands.

15 The ESWC, an acronym that first stood for "European Semantic Web Conference", was redefined to "Extended Semantic Web Conference" in 2010.

Table 12: Growth of publications since 2000 measured for the year 2009 for conference proceedings and journal publications

	Conference proceedings	Journal publications
BR	160%	188%
CA	85%	56%
CH	74%	49%
CN	1054%	318%
DE	70%	27%
DK	79%	39%
ES	230%	90%
FI	81%	30%
FR	89%	27%
GB	44%	21%
IN	404%	131%
IT	98%	59%
JP	48%	3%
KR	185%	173%
NL	57%	53%
RU	-8%	2%
SE	27%	27%
US	26%	25%
ZA	111%	98%
WORLD	141%	43%

Source: SCIE/CPCI; own calculations.

In 2007, a major positive deviation is observable for many countries, in particular Japan, South Korea, Spain, Switzerland, Finland and France. With the knowledge from the general coverage of the conference proceedings in the database this might be caused by the general increased coverage of proceedings in that particular year. Only Russia and Brazil show a slightly decreasing share in 2000 – 2009. Still, when comparing the absolute numbers in conference proceedings for these countries (Table 12), at least for Brazil the value is increasing. For Russia, the decreasing publication numbers are only true for conference proceedings, the number of journal publications are increasing – though not in a similar level as for most of the other countries. Again, China shows an exceptional growth rate both in conference as well as journal publications.

For most countries, the relative growth in conference proceeding numbers is higher than for journal publications. It is at this point hard to decide whether this is merely a database effect or really an increase in relevance of that publication format. The previous

results for the conferences (see Figure 10 and Table 8) suggest that even though there are few omissions of major conferences the growth in conference numbers and associated publications can be attributed to the inclusion of more smaller conferences. Again, it is hard to decide whether this is an effect caused by increased database coverage or not. However, conference trackers like http://www.allconferences.com seem to corroborate the notion that the number of conferences has increased tremendously in this time period. Still, such effects have to be born in mind when comparing growth rates of countries. We will look at the impact on the rankings of the countries in Section 4.2.

Table 13: Percentage of citations received after 5 years that was already emitted in a 3 year citation window

	All publications	Journal Publications	Conference Proceedings
BR	47%	47%	45%
CA	48%	48%	45%
CH	49%	49%	45%
CN	47%	47%	44%
DE	50%	50%	49%
DK	48%	48%	51%
ES	49%	49%	46%
FI	50%	50%	49%
FR	49%	49%	48%
GB	48%	48%	47%
IN	47%	47%	43%
IT	50%	50%	47%
JP	50%	50%	52%
KR	49%	49%	51%
NL	48%	48%	48%
RU	52%	52%	49%
SE	49%	49%	51%
US	48%	48%	46%
ZA	48%	48%	39%

Source: SCIE/CPCI; own calculations.

Despite the rather high shares of conference proceedings in publications only between 0.2 and 0.7 % of all citations emitted to publications of the country set in the year 2006 targeted conference proceedings. This observation holds for both a 3-year and a 5-year citation window. Still, we compared the citations received by the publications in these different time spans to answer the questions if conference proceedings need a

different citation window as journal publications or vice versa. In regard to the different time span between submission and publication and availability for the scientific community, one could assume that the timeliness of the citations might differ for the two publication types. In order to test this, we calculated the share of citations received in a 5-year-period that were already covered after 3 years. Table 13 shows the results for the year 2006 for our country set for the different publication types. Note that because of the small share of citations to conference proceedings, there is only a small difference between the share of citations for all publications and for journal publications. The prediction factor, i.e. the ratio between the citation values, of the 3-year citation window for the 5-year citation window is only slightly different for journal and conference publications. Thus, it cannot be derived that a 3-year citation window favors or penalizes any of these document types. Nonetheless, the major finding is that the inclusion of conference proceedings influences the overall citation number only slightly (Table 14). However, this also means that the citation rate is reduced more severely for countries with a higher share of conference proceedings.

Table 14: Citation rates for journal publications, conference proceedings and all documents in the year 2006

	All publications	Journal Publications	Conference Proceedings
BR	3.1	3.4	0.1
CA	5.4	6.1	0.1
CH	7.6	8.2	0.3
CN	2.6	3.6	0.0
DE	5.8	6.4	0.3
DK	7.0	7.4	0.2
ES	4.9	5.4	0.2
FI	5.6	6.3	0.1
FR	5.3	5.8	0.2
GB	5.9	6.4	0.2
IN	2.9	3.2	0.1
IT	5.2	5.9	0.2
JP	4.1	4.8	0.2
KR	3.1	3.8	0.1
NL	6.7	7.3	0.3
RU	2.4	2.7	0.1
SE	6.3	6.8	0.2
US	6.2	6.9	0.2
ZA	4.1	4.4	0.1

Source: SCIE/CPCI; own calculations.

Table 15: Citation rates in 27 scientific fields for journal publications, conference proceedings and all documents in the year 2006

Field	All publications	Journal Publications	Conference Proceedings
Basic Chemistry	5.5	5.6	0.3
Biology	5.3	5.3	0.5
Biotechnology	7.9	8.1	0.3
Chemical Engineering	2.8	3.0	0.2
Computers	1.0	2.3	0.1
Ecology, Climate	4.0	4.3	0.2
Electrical Engineering	1.3	3.5	0.1
Food, Nutrition	3.9	4.0	0.1
Geosciences	3.2	3.8	0.3
Humanities	0.5	0.5	0.1
Materials research	3.4	3.9	0.2
Mathematics	1.8	2.1	0.2
Measuring, Control	1.4	3.3	0.1
Mechanical engineering	1.4	2.2	0.1
Medical engineering	3.8	4.9	0.1
Medicine	5.5	5.6	0.2
Multidisciplinary	13.5	13.8	0.3
Nuclear technology	1.9	2.1	0.2
Optics	1.7	4.2	0.1
Organic Chemistry	5.6	5.6	0.1
Other	2.5	2.8	0.2
Pharmacy	5.7	5.8	0.1
Physics	4.3	5.2	0.3
Polymers	4.3	4.4	0.7
Social Sciences, Economics	1.7	2.3	0.0
Social Sciences, Other	2.4	2.6	0.1
Specific engineering	2.2	3.0	0.1

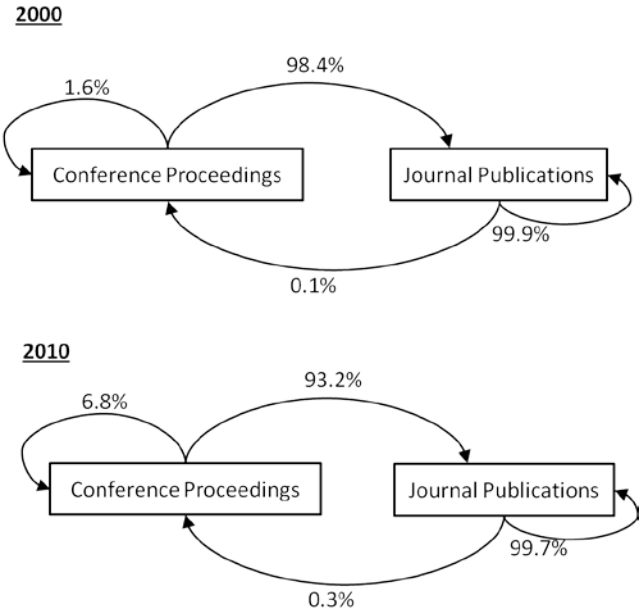
Source: SCIE/CPCI; own calculations.

For example, in the case of China and Great Britain, the citation number increases only by 0.5% and 0.3% with the inclusion of conference proceedings, while the publication number increases by 28% and 7% respectively in 2006 . Of course, the overall change in citation rates also depends on the citation rates of the document types. However, in this example we see that China's already comparatively low citation rate of 3.6 is re-

duced to 2.6 corresponding to a loss of 28%. This also results in a lower ranking according to the citation rates in our country set; out of the 19 countries, China is ranked 18 instead of 16. Great Britain on the other hand only loses 7% of its citation rate but even profits in the ranking by the inclusion of conference proceedings by gaining one rank (see Chapter 4.2 for a further discussion of these implications).

Table 15 shows the changes in citation rates for the fields with and without conference proceedings. The last column shows the decrease in the citation rate if conference proceedings are included. The citation rate for conference proceedings is smaller than that for journal publications in general so that the citation rate for the fields (or the countries as we could see in Table 14) can only suffer from the inclusion of conference proceedings. But the citation rate and thus the implicit acknowledgement of the fields drops of course especially for those fields that have a high percentage of conference proceedings. Even though small changes can be perceived (e.g. Computer Science or Mechanical Engineering), the order of Table 15 according to changes in citation rate mirrors the ranking of Figure 11 according to shares of conference proceedings to a large extent. Specific Engineering and Electrical Engineering have a comparable citation rate. However, Electrical Engineering has the second highest share of conference proceedings and thus also a higher decrease in citation rates.

Figure 12: Citation flows of conference proceedings and journal publications in 2000 and 2010 in percent



Source: Own illustration.

Meho and Yang (Meho/Yang 2007) stated that at least every 10th citation stems from conference papers. To test this for our dataset, we calculated the numbers and shares of document types in the references of the two kinds of documents. Figure 12 illustrates the citation flows from conference proceedings and articles for the years 2000 and 2010. For example, the arrow pointing from conference proceedings to articles in the upper figure shows that of all references in proceedings in the year 2000, 98.4% referred to articles. Note that the sum of the respective shares equals 100% even though we did not restrict the total reference count to these two document types specifically. Thus no other references are covered in the WoS. Despite the low coverage of the years at the margins of the analyzed time period the development depicted here is consistent over the whole time. References to conference proceedings are increasing in both document types over time. However, this also means the share of citations from conference proceedings to articles is decreasing. The slightly higher growth rate in references for conference proceedings leads to a reference number in 2009 that is 2.3 times as high as that in 2000. For comparison, the number of references for journal publications merely doubled in the same time period.

Table 16: Percentage of publications that have been republished as journal documents under the same title

Year	Resubmitted journal documents	Resubmitted conference proceedings
2000	0.41%	1.56%
2001	0.39%	1.47%
2002	0.44%	1.54%
2003	0.46%	1.32%
2004	0.47%	1.35%
2005	0.49%	1.16%
2006	0.50%	0.73%
2007	0.41%	0.62%
2008	0.27%	0.35%
2009	0.27%	0.18%
2010	0.11%	0.01%

Source: SCIE/CPCI; own calculations.

A further point of investigation is the number of resubmissions of both journal articles as well as conference proceedings. Table 16 shows the proportion of publications for which a later journal article with the same title and at least one shared author could be found. The percentage of republished conference proceedings is higher than that for journal documents until 2008. As was already noted, the remaining years seem to be relatively sparsely covered for conference proceedings so that the low numbers for

these years cannot be fully interpreted. Still, even though for the earlier years about 1.5% of the conference proceedings are republished a downward trend can be seen. This trend cannot be explained by the lower chance for resubmissions in later years because this does not affect the numbers for journal resubmissions in an equal way.

Table 17: Percentage of conference proceedings in resubmissions

first	**resubmission year**											
year	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**Total**
2000	14%	44%	48%	42%	39%	22%	21%	19%	13%	0%	4%	33%
2001		8%	46%	48%	45%	33%	31%	36%	11%	10%	19%	33%
2002			13%	39%	45%	43%	34%	30%	24%	6%	4%	32%
2003				10%	36%	47%	43%	33%	31%	13%	6%	30%
2004					11%	37%	44%	42%	34%	15%	5%	30%
2005						9%	35%	43%	44%	38%	27%	28%
2006							5%	26%	42%	39%	29%	21%
2007								6%	38%	43%	50%	25%
2008									10%	30%	50%	22%
2009										2%	34%	13%
2010											2%	2%
Total	14%	28%	33%	30%	30%	28%	26%	23%	29%	19%	24%	27%

Source: SCIE/CPCI; own calculations.

Table 17 provides another view on this issue; it depicts the percentage of journal articles that were formerly published as conference proceedings. The numbers on the diagonal represent the share that was republished in the same year. The calculations were made like the ones for the previous analysis. It can be seen that especially those resubmissions in a 2 to 4 years window are former conference proceedings publications. In total, every fourth resubmission stems from a former conference proceedings paper.

Table 18 shows the number of sources in the SCIE that had one of our keywords (see Chapter 2.2) in the title. They constitute approximately 1% of the publications in the SCIE. Also, it can be found that the number of publications per source each year is relatively high. So, we can derive that few but comparatively long conference proceedings are included in the SCIE.

More remarkably is that this source set seems to be rather fixed with a consistent ID assignment. In contrast to those proceedings covered in the CPCI, these sources take a stable proportion in the SCIE. Still, there exist sources that have multiple IDs. The "Institute of Physics Conference Series" for instance has 15 IDs in the time span 2000 – 2010.

Thus, a count of the number of distinct IDs of sources over the years 2000 – 2010 results in 107 entries in total, while the number of different source titles (without further cleaning) corresponds to 88 sources.

Table 18: Number of sources/publications in SCIE that are associated with a publication source title containing one of our keywords

Year	Sources	Publications	Share in SCIE
2000	66	10,396	1.19%
2001	64	9,932	1.14%
2002	64	9,797	1.10%
2003	70	10,483	1.14%
2004	68	10,734	1.11%
2005	68	11,297	1.12%
2006	62	10,462	0.98%
2007	54	9,357	0.84%
2008	57	9,174	0.77%
2009	58	10,267	0.82%
2010	55	9,753	0.76%

Source: SCIE; own calculations.

4.2 Implications for Bibliometric Analyses

In this section, we investigate the implications of the observations in the previous sections on bibliometric analyses. In particular, we take a look on various bibliometric indicators and the effects of the in- or exclusion of conference proceedings.

Table 19 shows the ranks of a selected set of countries in the year 2009 according to their share in the total publications covered in the database. Such rankings are oftentimes used to show how single countries contribute to the database or the scientific communication in general. For instance, the high share of the US does not only show the high presence in the scientific community, but also that effects noted in the database might be highly influenced by the US behavior. As can be seen in Table 19 the overall ranking is not affected by the exclusion of conference proceedings. The selected countries, which are the major contributors in the database, all but one profit from the exclusion of conference proceedings in terms of database shares. But most changes are on a minor scale. The big looser is China with nearly 3% of its share. As we have seen before, China has an extraordinarily high share of conference proceed-

ings in its publication set (cf. Table 11). The US, on the other hand, have a huge foundation of journal publications and an arbitrary share of conference proceedings.

Table 19: Shares and ranks of selected countries with in- or exclusion of conference proceedings within all publications in 2009

Country	Rank	Share with conference proceedings	Share without conference proceedings	Differences between without and with proceedings
US	1	25.80%	28.16%	2.36%
CN	2	13.15%	10.21%	-2.95%
GB	3	7.02%	7.83%	0.81%
DE	4	6.76%	7.20%	0.44%
JP	5	6.14%	6.26%	0.13%
FR	6	4.91%	5.28%	0.37%
CA	7	4.16%	4.54%	0.38%
IT	8	4.08%	4.30%	0.22%
ES	9	3.41%	3.66%	0.25%
IN	10	3.08%	3.29%	0.21%
KR	11	2.93%	3.06%	0.14%
BR	12	2.32%	2.58%	0.26%
NL	13	2.25%	2.49%	0.25%
RU	14	2.03%	2.29%	0.25%
CH	15	1.60%	1.79%	0.19%
SE	16	1.43%	1.60%	0.17%
DK	17	0.81%	0.92%	0.10%
FI	18	0.77%	0.81%	0.04%
ZA	19	0.56%	0.63%	0.07%

Source: SCIE/CPCI; own calculations.

Table 20: Ranks of selected countries according to citation rate with in- or exclusion of conference proceedings in 2006

Country	Including conference proceedings	Excluding conference proceedings
CH	1	1
DK	2	2
NL	3	3
SE	4	5
US	5	4
GB	6	7
DE	7	6
FI	8	8
CA	9	9
FR	10	11
IT	11	10
ES	12	12
JP	13	13
ZA	14	14
KR	15	15
BR	16	17
IN	17	18
CN	18	16
RU	19	19

Source: SCIE/CPCI; own calculations.

Table 21: Citation rate of selected countries for different document type combinations in 2006 (3 year citation window)

Country	Citation rate with conference proceedings	Citation rate without conference proceedings	Citation rate conference proceedings only
BR	3.13	3.45	0.12
CA	5.44	6.13	0.15
CH	7.59	8.16	0.31
CN	2.58	3.58	0.04
DE	5.84	6.41	0.28
DK	6.97	7.40	0.19
ES	4.89	5.37	0.15
FI	5.56	6.28	0.13
FR	5.27	5.84	0.20
GB	5.91	6.36	0.23
IN	2.85	3.18	0.06
IT	5.20	5.87	0.17
JP	4.11	4.76	0.19
KR	3.14	3.81	0.12
NL	6.75	7.34	0.26
RU	2.42	2.67	0.13
SE	6.31	6.81	0.19
US	6.18	6.88	0.22
ZA	4.10	4.40	0.13

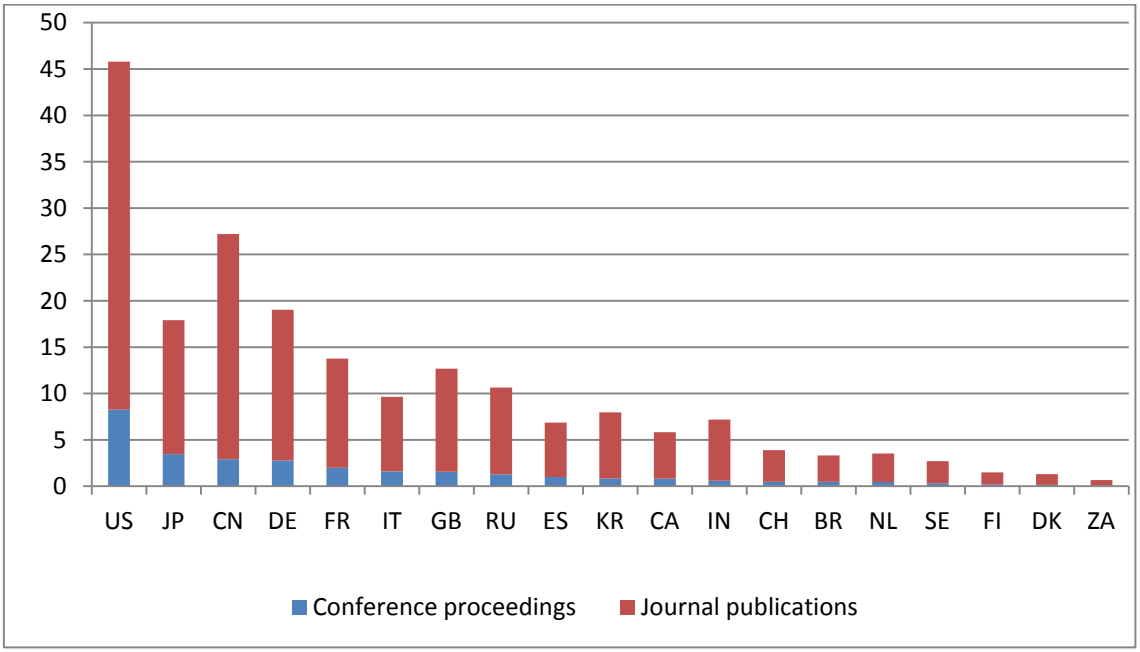
Source: SCIE/CPCI; own calculations.

However, the implications for the citation rate vary, as we have already seen in the descriptive analysis (Table 14 and Table 15). Table 20 shows the ranking of the countries regarding their citation rate. In this case as well, the changes due to the exclusion of the conference proceedings are of minor nature. China is penalized for its high share in conference proceedings with a loss in ranks of 2 because of the lower citation rate of these documents. Therefore, we can only derive that the use of conference proceedings can have a positive or negative effect for those countries with a high proceedings share. But as Table 21 shows, China also has an extraordinarily low citation rate for the conference proceedings.

The results so far suggest that because of the quite distinct citation rates for both document types, analyses that are supposed to consider both document types better do this in separate reportings. A mixture of both document types in a bibliometric analysis that at least includes citation rates could be skewed because of the lower citation rate for conference proceedings. The interpretation of the citation rate then becomes mere guessing – is a low/high citation rate caused by high/low shares of conference proceeding or are there any other reasons? The mixture of two document types with such

a difference in average citation rate makes any valid interpretation difficult. Thus, either a separate analysis or a normalization according to the document type (e.g. the average citation rate in that field for that document type) pose a better option.

Figure 13: Number of publications (in thousands) in Physics with in- or exclusion of conference proceedings in 2009



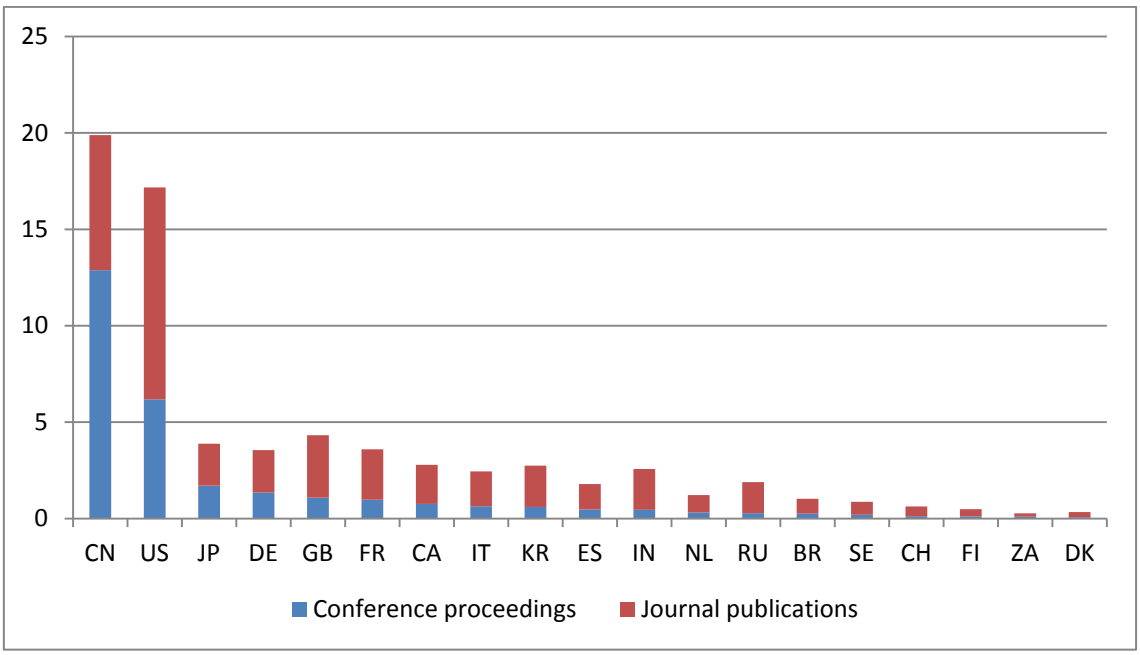
Source: SCIE/CPCI; own calculations and illustration.

As we saw in the previous analysis, conference proceedings have different levels of significance and thus different shares in scientific fields (cf. Figure 11). Therefore, if the overall shares of conference proceedings of analyzed subjects are equal, they might still be concerned with different fields that demand different levels of scientific output. According to Figure 11, Physics, Electrical Engineering, Computer Science, Mechanical Engineering, Optics and Measuring are the fields that are most affected by the conference proceedings. We will therefore analyze these fields specifically in the following.

Figure 13 to Figure 15 show the absolute number of publications in the fields Physics, Mechanical Engineering and Computer Science. The countries are ordered on the x-axis according to their number of conference proceedings (in blue). China ranks first twice and third once in these fields corroborating our previous results – the high percentage of conference proceedings seems to stem from a high specialization on these particular fields.[16]

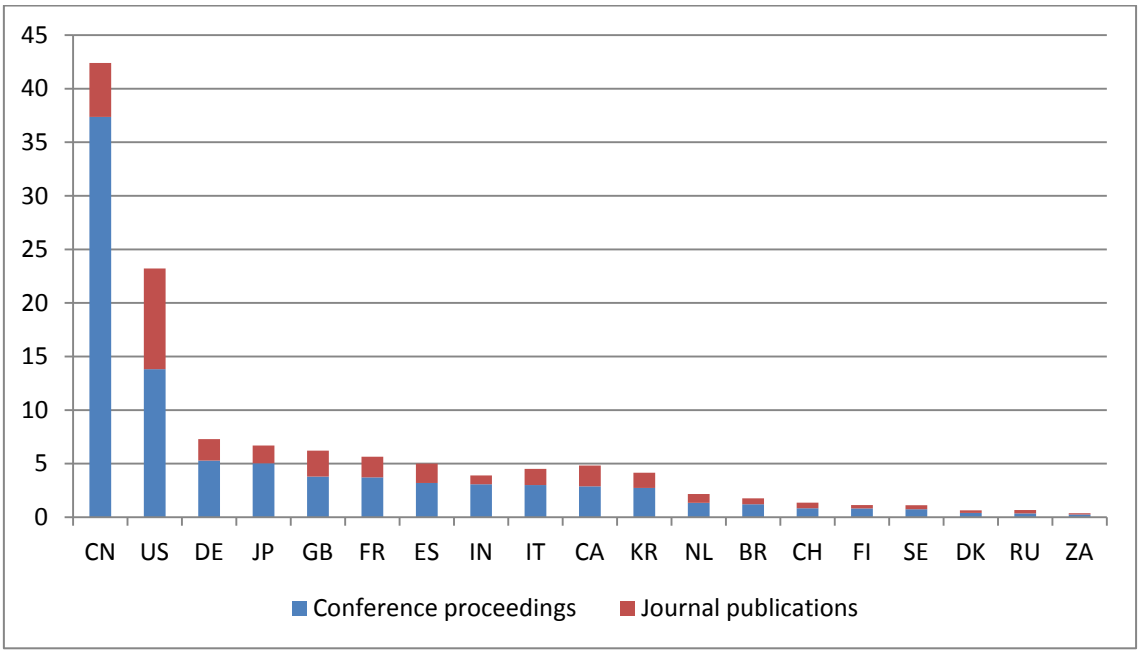
16 This notion is also confirmed later in Table 22.

Figure 14: Number of publications (in thousands) in Mechanical engineering with in- or exclusion of conference proceedings in 2009



Source: SCIE/CPCI; own calculations and illustration.

Figure 15: Number of publications (in thousands) in Computer Science with in- or exclusion of conference proceedings in 2009



Source: SCIE/CPCI; own calculations and illustration.

The US rank second twice and first once showing a high representation in these fields as well. In Mechanical Engineering though, the US rank second behind China both in

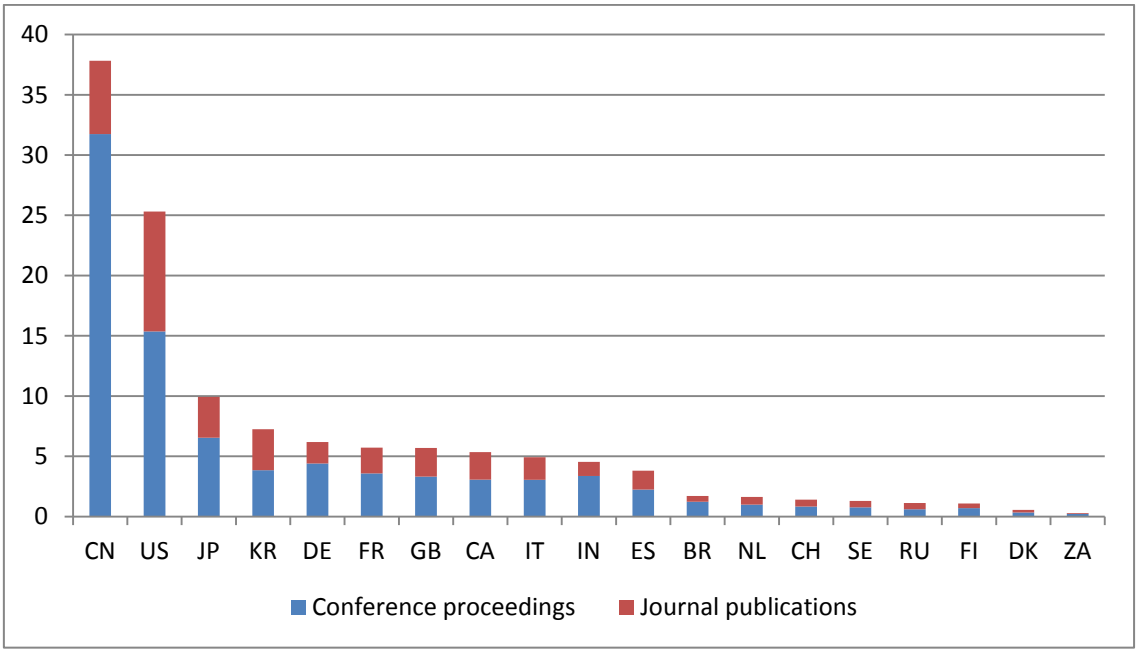
total and in conference proceedings numbers, so one would deduce that China has a bigger focus/specialization here. However, if the analysis was reduced to journal publications only, the US would lead (see Table 22). This suggests that the US are a more important key player in Mechanical Engineering. The results also indicate that at least in some fields, the publication output in conference proceedings is not necessarily mirrored in journal publications, e.g. China leads in Computer Science in conference proceedings, but journal publications represent only a small portion of its research. In Physics this effect is not notable, i.e. the rankings stay the same whether proceedings are included or not but here we have a quite contrary effect: According to Figure 13, Japan ranks second regarding conference proceedings. Using the numbers for journal publications or journal publications and conference proceedings, it only ranks 4th. Thus, Japan, which basically mirrors the share of conference proceedings in Physics in general with approximately 20% in overall publications (see Figure 13), is penalized in rankings if journal publications are included.

Table 22: Ranks of selected countries with in- or exclusion of conference proceedings in the fields Physics, Engineering and Computer Science in 2009

	Physics		Mechanical Engineering		Computer Science	
Country	Rank with conference proceedings	Rank without conference proceedings	Rank with conference proceedings	Rank without conference proceedings	Rank with conference proceedings	Rank without conference proceedings
US	1	1	2	1	2	1
CN	2	2	1	2	1	2
DE	3	3	6	5	3	4
JP	4	4	4	6	4	8
FR	5	5	5	4	6	6
GB	6	6	3	3	5	3
RU	7	7	11	11	17	17
IT	8	8	10	10	9	9
KR	9	9	8	7	10	10
IN	10	10	9	8	11	12
ES	11	11	12	12	7	7
CA	12	12	7	9	8	5
CH	13	13	16	16	14	14
NL	14	14	13	13	12	11
BR	15	15	14	14	13	13
SE	16	16	15	15	16	15
FI	17	17	17	17	15	16
DK	18	18	18	18	18	18
ZA	19	19	19	19	19	19

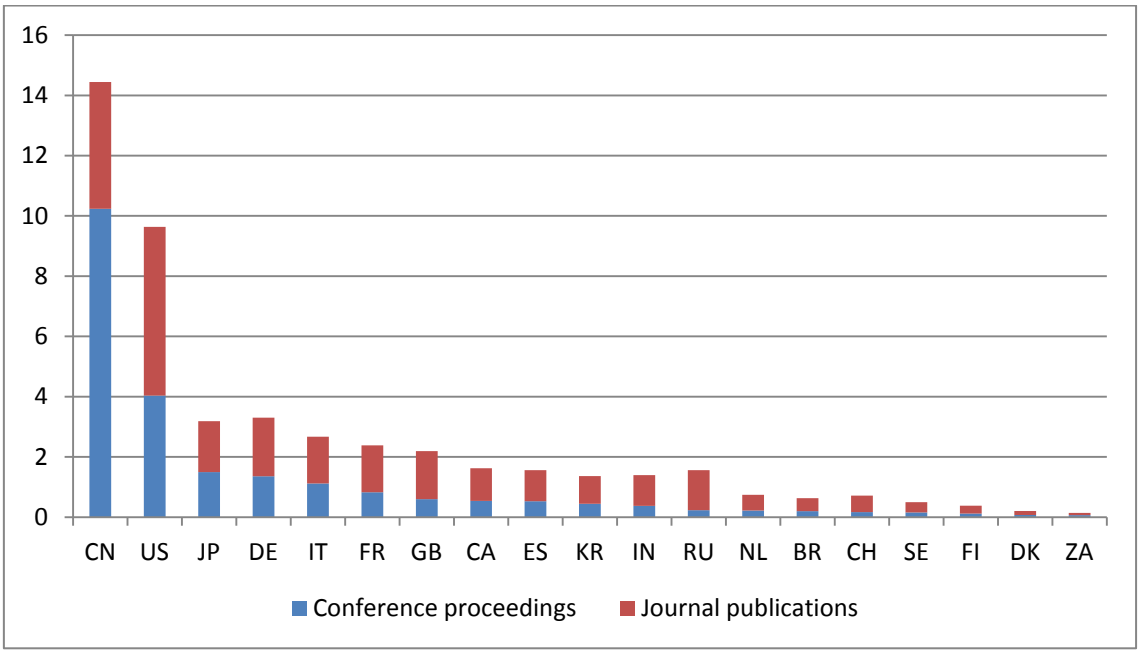
Source: SCIE/CPCI; own calculations.

Figure 16: Number of publications (in thousands) in Electrical engineering with in- or exclusion of conference proceedings in 2009



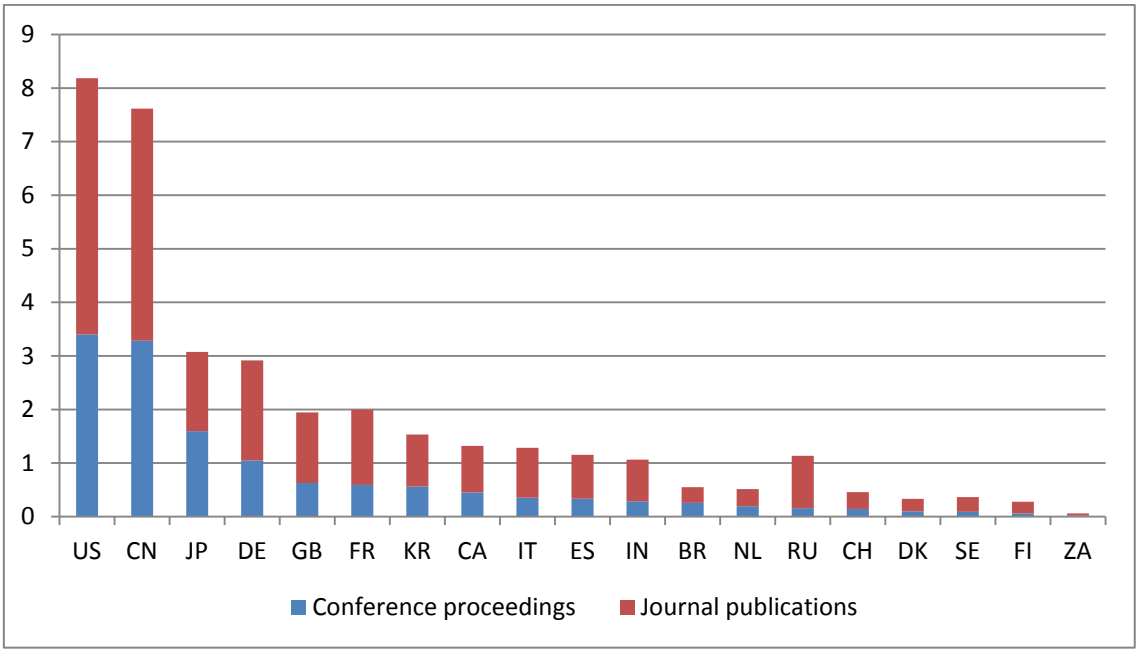
Source: SCIE/CPCI; own calculations and illustration.

Figure 17: Number of publications (in thousands) in Measuring, control with in- or exclusion of conference proceedings in 2009



Source: SCIE/CPCI; own calculations and illustration.

Figure 18: Number of publications (in thousands) in Optics with in- or exclusion of conference proceedings in 2009



Source: SCIE/CPCI; own calculations and illustration.

In Mechanical Engineering as well as Computer Science, changes in the rankings due to in- or exclusion of conference proceedings nearly affect the whole scale and lead to differences of up to 4 positions (Table 22, Japan in Computer Science).

Figure 16 to Figure 18 show the similar analysis for the remaining fields, which heavily rely on conference proceedings: Electrical Engineering, Measuring, Control and Optics. Again, the US and China share the two first places according to conference proceedings. In Measuring, Control and Optics Russia deviates with a relatively high journal publication output that would lead to a different ranking if journal publications were added to the analysis.

Again, similar results can be found for the comparison of the rankings according to the number of publications (see Table 23). In the fields Electrical Engineering and Measuring, Control China and the US switch positions when conference proceedings are included because China's conference proceedings output clearly outranks that of the US. In Electrical Engineering, the ranks up to position 17 are mixed up when conference proceedings are included leading to changes up to 4 ranks (Germany and Brazil). Russia profits in all cases from the inclusion of conference proceedings.

Regarding these results, it is difficult to provide a general rule to decide whether conference proceedings should be used when comparing the publication output of the countries. As we already saw in Table 19, the field specific effects seem to more or less cancel each other out when they are summarized. Thus, the decision for or against

conference proceedings does not affect the overall results. However, in particular for analyses of single fields, we suggest a decision that depends on the standing of conference proceedings in the specific field. For example, the fields shown above have according to Figure 11 a high focus on conference proceedings. A lot of the dissemination in these fields happens at conferences. Thus, conference proceedings should be analyzed for these fields as well.

To see how the portfolio of the single countries was affected by the different document type choices, we calculated the RLA Index as defined by Hinze and Grupp (1996). The RLA Index denotes the specialization of a single country in a single field in comparison to the world average. Thus, if a larger/smaller share of a country *c*'s publication belong to a field *f* than is the case worldwide, a positive/negative value is assigned to the RLA Index for *f* and *c*. The value lies within the range of -100 and 100 and the absolute value denotes the extent of the deviation from the world average.

Table 23: Ranks of selected countries with in- or exclusion of conference proceedings in the fields Electrical engineering, Measuring, Control and Optics in 2009

Country	Electrical Engineering		Measuring, Control		Optics	
	Rank with conference proceedings	Rank without conference proceedings	Rank with conference proceedings	Rank without conference proceedings	Rank with conference proceedings	Rank without conference proceedings
CN	1	2	1	2	2	2
US	2	1	2	1	1	1
JP	3	4	4	4	3	4
KR	4	3	12	12	7	8
DE	5	9	3	3	4	3
FR	6	7	6	6	5	5
GB	7	5	7	5	6	6
CA	8	6	8	9	8	10
IT	9	8	5	7	9	9
IN	10	11	11	11	12	12
ES	11	10	9	10	10	11
BR	12	16	15	15	13	15
NL	13	12	13	14	14	13
CH	14	13	14	13	15	14
SE	15	14	16	16	16	16
RU	16	15	10	8	11	7
FI	17	17	17	17	18	18
DK	18	18	18	18	17	17
ZA	19	19	19	19	19	19

Source: SCIE/CPCI; own calculations.

Table 24: Changes in Germany's specialization (RLA Index) with in- or exclusion of conference proceedings in 2009

Field	RLA with conference proceedings	RLA without conference proceedings	Difference between in- and excluding proceedings
Basic chemistry	6.96	-0.18	7.14
Organic chemistry	7.72	1.50	6.22
Medicine	9.12	3.17	5.95
Polymers	-2.65	-8.59	5.94
Biology	6.66	0.80	5.86
Other	-28.16	-33.96	5.80
Pharmacy	-10.20	-15.99	5.79
Chemical engineering	-27.29	-32.94	5.66
Food, nutrition	-26.54	-32.09	5.55
Multidisciplinary	-2.81	-7.58	4.76
Humanities	-53.23	-57.93	4.71
Biotechnology	16.21	12.86	3.35
Materials research	1.68	-1.56	3.23
Electrical engineering	-48.78	-51.33	2.55
Geosciences	21.63	20.05	1.58
Ecology, climate	-6.70	-8.20	1.50
Nuclear technology	53.70	52.53	1.16
Physics	38.25	38.23	0.01
Mechanical engineering	-45.26	-45.11	-0.15
Social Sciences, Other	-35.92	-35.45	-0.47
Optics	10.41	13.53	-3.12
Medical engineering	23.57	27.98	-4.40
Specific engineering	-42.47	-36.92	-5.55
Mathematics	-22.21	-14.19	-8.02
Computer Science	-34.32	-25.61	-8.71
Measuring, control	-12.12	2.33	-14.45
Social Sciences, Economics	-33.88	-13.03	-20.85

Source: SCIE/CPCI; own calculations.

As a first example, Table 24 shows the RLA Index for Germany in the year 2009. Changes in Germany's specialization happen on different scales but most notably, there are also some cases in which an implicit specialization in one field is negated or vice versa. For instance, while German engineers seem to have less publications in journals and at conferences in Measuring, control, they have slightly more journal publications than the world average in this field. This leads to a different interpretation in the end. Similar observations can be made for Chemistry (Basic and Organic). These

findings suggest a different publication behaviour in Germany in these fields than in other countries.

Table 25: Changes in China's specialization (RLA Index) with in- or exclusion of conference proceedings in 2009

Field	RLA with conference proceedings	RLA without conference proceedings	Difference between in- and excluding proceedings
Social Sciences, Economics	45.39	-72.70	118.10
Social Sciences, Other	-60.67	-92.88	32.22
Computer Science	62.84	30.62	32.22
Medical engineering	-24.89	-49.56	24.67
Electrical engineering	54.51	29.98	24.53
Mechanical engineering	51.50	31.18	20.33
Measuring, control	59.67	41.83	17.84
Specific engineering	47.69	33.07	14.62
Mathematics	55.25	44.84	10.41
Humanities	-91.97	-94.72	2.74
Ecology, climate	-8.51	-2.70	-5.81
Geosciences	-9.45	1.16	-10.61
Medicine	-75.94	-65.05	-10.89
Polymers	55.58	68.20	-12.62
Other	-39.19	-26.22	-12.97
Materials research	49.46	65.46	-16.00
Food, nutrition	-56.31	-38.76	-17.54
Optics	37.93	55.63	-17.70
Basic chemistry	40.07	59.63	-19.55
Nuclear technology	-67.80	-46.96	-20.84
Biology	-53.32	-32.35	-20.96
Biotechnology	-20.02	1.07	-21.09
Pharmacy	-24.92	-3.33	-21.59
Chemical engineering	16.56	41.36	-24.79
Organic chemistry	4.41	29.24	-24.83
Multidisciplinary	-16.26	11.79	-28.05
Physics	9.44	42.74	-33.30

Source: SCIE/CPCI; own calculations.

For China, similar observations can be made for other fields (Table 25). In this case most notably the field Economics, which is the one with the highest specialization when conference proceedings are included, shows a very low negative value for journal publications only. According to Figure 11, conference proceedings normally account for 30 % of the publications in this field. China seems to tremendously exceed this value. It is

hard to judge whether this represents a real specialization or a drift to minor quality publications since the acceptance threshold for conference proceedings might be lower. However, leaving possible interpretations and reasons aside, the point is that this huge gap has to be addressed and noted when dealing with such analyses. Similar observations – in a smaller scale - can be made for the fields Geosciences and Multidisciplinary. And even when a change of sign is not notable, shifts in terms of specialization level make a comparison with other countries even more difficult. For example, the specialization level of Germany in Physics is not affected by a change in the document set while China is either less (with conference proceedings) or more specialized (without conference proceedings) than Germany.

Of the five countries that we selected for the field specialization analysis, only China shows a specialization in Measuring, control in both document sets. Japan (Table 26) for instance loses it's specialization not only in Measuring, control but also in Medical and Electrical Engineering when conference proceedings are included. On the other hand, the publication output in Optics seems to be favored by an analysis with conference proceedings. However, most of the fields in Japan are unaffected by the inclusion of conference proceedings. It becomes in particular obvious with the exemplary comparison of the results of Japan and Germany, that the national policy and handling of conference proceedings as well as maybe the propagation of conferences in the vicinity are important influencing factors.

Table 26: Changes in Japan's specialization (RLA Index) with in- or exclusion of conference proceedings in 2009

Field	RLA with conference proceedings	RLA without conference proceedings	Difference between in- and excluding proceedings
Optics	24.85	4.08	20.78
Specific engineering	-8.68	-20.57	11.89
Mechanical engineering	-29.15	-34.24	5.09
Medicine	-0.47	-2.86	2.39
Materials research	33.28	30.95	2.33
Ecology, climate	-32.20	-34.36	2.16
Biology	-4.81	-6.78	1.97
Basic chemistry	22.78	21.28	1.49
Organic chemistry	51.70	50.47	1.23
Pharmacy	27.98	26.78	1.20
Food, nutrition	22.97	21.85	1.11
Physics	41.31	40.22	1.09
Geosciences	-13.34	-14.38	1.04
Chemical engineering	-23.95	-24.96	1.00
Humanities	-97.02	-97.43	0.40
Polymers	46.60	46.38	0.21
Biotechnology	26.20	26.09	0.11
Multidisciplinary	-41.32	-41.22	-0.10
Social Sciences, Other	-90.35	-90.10	-0.25
Other	0.26	1.24	-0.98
Nuclear technology	57.90	60.17	-2.26
Computer Science	-33.27	-29.84	-3.43
Mathematics	-47.77	-43.73	-4.04
Social Sciences, Economics	-86.86	-79.45	-7.41
Measuring, control	-6.08	2.21	-8.28
Medical engineering	1.72	10.38	-8.66
Electrical engineering	3.63	20.67	-17.03

Source: SCIE/CPCI; own calculations.

For South Korea as well the changes in most fields are of minor, negligible nature. Still, especially in the Engineering and Computer Sciences (bottom five entries in Table A1 in the annex) and also in Economics notable shifts can be perceived. Since the normal definition of the RLA Index only includes journal publication, a high level of specialization for South Korea is suggested. Nevertheless, at least for Measuring, Control and Computer Science the level of dissemination and exchange with other researchers via conferences is below the worldwide average. Similar results can be seen for Finland (Table A2 in the annex). For all five of the selected countries, the so far often men-

tioned and conference proceedings focused fields Medical, Mechanical and Electrical Engineering as well as Measuring, Control, Optics and Economics are affected the most by the exclusion of conference proceedings. Only in the case of China, this effect is positive for the specific fields.

5 Conclusion

This study aimed at giving an overview over the coverage of conference proceedings and implications for appropriate usage of this kind of documents in bibliometric analyses.

Some of the findings suggest that conference proceedings and journal publications should not be used together when performing bibliometric analyses. There are many factors in which the characteristics of both data types differ. First, the coverage of conference proceedings seems to be less for more recent years. A combination of both document types would thus lead to an advantage for countries and fields respectively with a broader coverage in journal articles. Second, the citation behaviour for (and of) conference proceedings differs. Conference proceedings are only cited by a small fraction of the publications in the database and therefore have a relatively low citation rate in comparison with journal articles. It was discussed that the share of conference proceedings influences the overall citation rate. Third, the coverage and availability of sources is more volatile for conference proceedings. As we saw both in the largest conferences in the database as well as the most important conferences in the fields, there are gaps that lead to an unreliable data set.

As Section 4.1 showed, the implications for the results of the bibliometric analysis are manifold. Even though the overall shares of the single countries in the database change only slightly, as soon as single fields are involved, the different standing of conference proceedings in the fields influences the outcome. Especially fields in which publishing at conferences is an important aspect can have quite different interpretations with and without conference proceedings. The results indicate that an additional analysis of or with conference proceedings can provide further hints.

Finally, as Section 3 showed, exemplary countries, the US, Germany, Great Britain, Japan and China, display diverse trends of bibliometric indicators as well as according ranks for journal publications and conference proceedings, implying two kinds of publication activities for the two kinds of publication types, and the necessity of respective bibliometric analyses.

Possible future work would be a broader analysis of covered and not covered conferences and a comparison with the also volatile coverage of journals. Also, the results for the RLA Index suggest that conference proceedings are deemed more valuable in some countries than in others. An in depth analysis of this aspect for instance supported by a qualitative analysis could help to better understand the mechanisms behind these findings.

References

Abraham, P. (2000): Duplicate and salami publications, *Journal of Postgraduate Medicine,* 46, 67-69.

Aleixandre-Benavent, R./González-Alcaide, G./Miguel-Dasit, A./Navarro-Molina, C./Valderrama-Zurián, J.C. (2009): Full-text publications in peer-reviewed journals derived from presentations at three ISSI conferences, *Scientometrics,* 80, 407-418.

Bar-Ilan, J. (2010): Web of Science with the Conference Proceedings Citation Indexes: the case of computer science, *Scientometrics,* 83, 809-824.

Cocosila, M./Serenko, A./Turel, O. (2011): Exploring the management information systems discipline: A scientometric study of ICIS, PACIS and ASAC, *Scientometrics,* 87, 1-16.

Drott, M.C. (1995): Reexamining the Role of Conference papers in Scholarly Communication, *Journal of the American Society for Information and Technology,* 46, 299-305.

Eckmann, M./Rocha, A./Wainer, J. (2012): Relationship between high-quality journals and conferences in computer vision, *Scientometrics,* 90, 617-630.

Franceschet, M. (2010): The skewness of computer science, *Information Processing and Management,* 47, 117-124.

Glänzel, W./Schlemmer, B./Schubert, A./Thijs, B. (2006): Proceedings literature as additional data source for bibliometric analysis, *Scientometrics,* 68, 457-473.

Harzing, A.-W. (2013): Document categories in the ISI Web of Knowledge: Misunderstanding the Social Sciences?, *Scientometrics,* 94, 23-43.

Hinze, S./Grupp, H. (1996): Mapping of R&D structures in transdisciplinary areas: New biotechnology in food sciences, *Scientometrics,* 37, 313-335.

Hofer, K.M./Smejkal, A.E./Bilgin, F.Z./Wuehrer, G.A. (2010): Conference proceedings as a matter of bibliometric studies: The Academy of International Business 2006-2008, *Scientometrics,* 84, 845-862.

Íñiguez-Rueda, L./Martínez-Martínez, L.M./Muñoz-Justicia, J./Peñaranda-Cólera, M.C./Sahagún-Padilla, M.Á./Alvarado, J.G. (2008): The mapping of Spanish social psychology through its conferences: A bibliometric perspective, *Spanish Journal of Psychology,* 11, 137-158.

Kademani, B.S./Sagar, A./Kumar, V. (2009): Conference papers of BARC scientists and engineers: A citation based study, Proceedings of the National Conference on "Putting Knowledge to work: Best Practices in Librarianship". Mumbai: Scientific Information Resource Division, Knowledge Management Group, Bhabha Atomic Research.

Lisée, C./Larivière, V./Archambault, É. (2008): Conference proceedings as a source of scientific information: A bibliometric analysis, *Journal of the American Society for Information Science and Technology,* 59, 1776-1784.

Lv, P.H./Wang, G.F./Wan, Y./Liu, J./Liu, Q./Ma, F.C. (2011): Bibliometric trend analysis on global graphene research, *Scientometrics,* 88, 399-419.

Meho, L.I./Rogers, Y. (2008): Citation counting, citation ranking, and h-Index of Human-Computer Interaction Researchers: A Comparison between Scopus and Web of Science, *Journal of the American Society for Information and Technology,* 59, 1711-1726.

Meho, L.I./Yang, K. (2007): Impact of data sources on citation counts and rankings of LIS faculty: Web of science versus scopus and google scholar, *Journal of the American Society for Information and Technology,* 58, 2105-2125.

Moed, H.F./Visser, M.S. (2007): *Developing Bibliometric Indicators of Research Performance in Computer Science: An Exploratory Study*, Research Report to the Council for Physical Sciences of the Netherlands Organisation for Scientific Research. The Hague: Council for Physical Sciences of the Netherlands Organisation for Scientific Research.

Montesi, M./Owen, J.M. (2008): From conference to journal publication: How conference papers in software engineering are extended for publication in journals, *Journal of the American Society for Information and Technology,* 59, 816-829.

Sigogneau, A. (2000): An analysis of document types published in journals related to physics: Proceeding papers recorded in the Science Citation Index database, *Scientometrics,* 47, 589-604.

Copyright

Annex

Table A1: Changes in South Korea's specialization (RLA Index) with in- or exclusion of conference proceedings in 2009

Fields	RLA with conference proceedings	RLA without conference proceedings	Difference between in- and excluding proceedings
Organic chemistry	-27.65	-31.47	3.82
Medicine	-20.97	-24.63	3.66
Pharmacy	30.35	26.85	3.50
Biology	-36.21	-39.61	3.40
Basic chemistry	28.97	26.16	2.81
Multidisciplinary	-76.87	-79.45	2.58
Food, nutrition	45.23	43.02	2.21
Biotechnology	15.37	13.28	2.08
Other	-39.94	-40.64	0.70
Humanities	-92.75	-93.43	0.68
Chemical engineering	37.59	37.41	0.18
Ecology, climate	-36.09	-35.55	-0.54
Nuclear technology	51.71	52.25	-0.54
Social Sciences, Other	-81.15	-80.59	-0.56
Polymers	60.55	61.19	-0.64
Geosciences	-63.61	-62.31	-1.30
Materials research	53.57	55.43	-1.86
Optics	28.89	32.21	-3.32
Physics	35.32	40.96	-5.64
Specific engineering	7.18	17.91	-10.72
Mathematics	-40.28	-28.13	-12.15
Social Sciences, Economics	-66.98	-50.40	-16.58
Medical engineering	19.84	38.21	-18.37
Mechanical engineering	9.12	32.81	-23.69
Measuring, control	-16.88	12.19	-29.07
Electrical engineering	43.09	72.98	-29.89
Computer Science	-8.28	24.24	-32.51

Source: SCIE/CPCI; own calculations.

Table A2: Changes in Finland's specialization (RLA Index) with in- or exclusion of conference proceedings in 2009

Fields	RLA with conference proceedings	RLA without conference proceedings	Difference between in- and excluding proceedings
Chemical engineering	-8.85	-27.85	19.01
Organic chemistry	-24.01	-30.15	6.14
Humanities	-43.52	-48.95	5.43
Medicine	10.27	5.29	4.98
Basic chemistry	-28.26	-32.04	3.77
Biology	24.29	20.54	3.75
Pharmacy	-23.99	-27.36	3.37
Food, nutrition	37.19	33.92	3.26
Materials research	-33.19	-35.99	2.79
Other	33.00	30.78	2.22
Multidisciplinary	-29.04	-31.19	2.15
Social Sciences, Other	7.85	6.13	1.71
Biotechnology	18.48	16.79	1.69
Specific engineering	-38.39	-38.70	0.31
Polymers	-37.71	-37.90	0.19
Geosciences	13.47	13.79	-0.32
Nuclear technology	42.02	44.56	-2.54
Ecology, climate	46.34	48.90	-2.56
Physics	2.96	7.11	-4.16
Mathematics	-28.24	-22.69	-5.54
Medical engineering	-8.03	0.77	-8.81
Computer Science	-4.97	8.11	-13.08
Social Sciences, Economics	23.65	39.86	-16.20
Electrical engineering	-10.59	8.54	-19.13
Mechanical engineering	-29.80	-8.54	-21.26
Optics	-7.47	18.24	-25.71
Measuring, control	-11.00	20.42	-31.43

Source: SCIE/CPCI; own calculations.