Title: A citation analysis of the research reports of the Central Mining Institute: a case study: mining and environment using the Web of Science, Scopus, BazTech, and Google Scholar

Author: Magdalena Bemke-Świtilnik, Aneta Drabek

A citation analysis of the research reports of the Central Mining Institute. Mining and Environment using the Web of Science, Scopus, BazTech, and Google Scholar: A case study

Magdalena Bemke-Świtilnik*, Aneta Drabek

Training and Information Centre, Central Mining Institute, Plac Gwarków 1, 40-166 Katowice, *Library of University of Silesia, Bankowa 11a, 40-007 Katowice, Poland

ABSTRACT

This paper presents the analysis of a Polish mining sciences journal (Prace Naukowe GIG. Górnictwo i Środowisko; title in English: Research Reports of the Central Mining Institute. Mining and Environment; acronym in English [RRCMIME]). The analysis is based on data from the following sources: the Web of Science (WoS), Scopus, BazTech (a bibliographic database containing citations from Polish Technical Journals), and Google Scholar (GS). The data from the WoS and Scopus were collected manually as RRCMIME is not indexed in these sources. GS does index BazTech; however, research shows that only 84% of BazTech citations are available through GS. Therefore, this may suggest that the most recommended sources for the bibliometric analyses of domestic journals, which are not indexed in international databases, are national data sources.

Keywords: BazTech, citation analysis, domestic journal, Google Scholar, Scopus, Web of Science

INTRODUCTION

Currently, there are about 28 Polish domestic journals in the field of mining sciences and engineering geology. They are primarily published by Polish mining science institutions such as technical universities, research institutes, units of the Polish Academy of Sciences, and other specialized units (e.g., office of government administration such as the State Mining Authority, or specialized publishing houses). These journals are not indexed in the WoS™ Core Collection (WoS) or, with two exceptions, in Scopus. In this article, we present a case study of a Polish domestic journal in mining sciences, Prace Naukowe GIG. Górnictwo i Środowisko (title in English: Research Reports of the Central Mining Institute. Mining and Environment; acronym in English [RRCMIME]; ISSN 1643-7608). The citation window is 11 years (2002–2012), which covers the full span of RRCMIME publications. Since 2013, the journal has been publishing under the new title, Journal of Sustainable Mining, as the editors adopted a new development strategy.

So far, bibliometric analyses of RRCMIME have not been carried out. The aim of this study was to determine the publication activity and impact of the journal as well as to present a citation analysis. A few data sources were selected for the study: The WoS, Scopus, Google Scholar (GS), and the BazTech database which index the...
contents of Polish Technical Journals. Multiple data sources were assessed to discover their applicability when it comes to analyzing Polish Journals in Mining Sciences. This is the first time that BazTech was used in bibliometric studies.

LITERATURE REVIEW

Over the years, numerous bibliometric analyses of individual journals have been carried out. There are two review papers dealing with this matter. Tiew reviewed 102 publications from 1969 to 1997. Tiew’s work was continued by Wan et al. These authors included 82 cases of bibliometric analyses of individual journals from 1998 to 2008.

Citation analyses are typically carried out for journals indexed in large, multidisciplinary databases such as WoS or Scopus. The databases also enable the collection of citations from journals that are not indexed in these sources. However, data collection demands the adoption of a more time- and work-consuming strategy than for journals indexed in WoS or Scopus. There are publications that concern bibliographic studies on individual journals which are not indexed in WoS; however, the authors used this source in their research. With the subject matter, two similar citation analyses of Polish journals were found. Racki carried out a citation analysis of Geological Quarterly and presented the results in multiple contexts, highlighting the international impact of the journal. It is worth mentioning that the author did a number of other bibliometric studies on Polish journals concerning the sciences as well.

Another example of citation analysis of individual journals not indexed in WoS is the study of two Korean Medical Journals namely the Korean Journal of Medicine and the Korean Journal of Urology. Huh used several sources such as WoS, Scopus, Gs, and National Information Sources – Korean Medical Citation Index and KoreaMed Synapse. In doing so, it became possible to see the differences in the number of citations of the journal depending on the data source.

Citation analyses of special issues of journals that contain proceedings papers (PPJs) are an area that must be focused on. PPJs are a significant part of the scholarly communication process. An advantage of PPJ is its use of a two-stage process for reviewing. First, the author receives comments from conference attendees and then from peer reviewers. The citation analysis of a special issue of industrial health showed that PPJ was at least 3 times more cited than a regular issue and had a higher immediacy index. Studies on a larger scale in science and social sciences prove, however, that PPJ gets fewer citations than regular issues.

This study has been carried out to answer the following research questions:

- Is RRCMIME as a domestic journal cited by journals indexed in WoS or Scopus?
- What is the share of citations of RRCMIME in International Journals compared with domestic journals?
- Do RRCMIME papers receive citations from authors outside of Poland?
- Does the database that indexes Polish journals give higher bibliometric indicators (total number of citations, average number of citations per paper, journal self-citations, and H-index) for RRCMIME than for other sources?
- What is the share of PPJ in the citation impact of RRCMIME?
- How soon are RRCMIME papers cited?
- Which database is the most representative for the citation analysis of selected technical domestic journals?

DATA AND METHODS

To find citations, all data concerning all the papers published in RRCMIME were used. The journal was published between the years 2002 and 2012 in printed form. In 2013, RRCMIME changed its title to Journal of Sustainable Mining. Because of the short period of publishing, papers from the Journal of Sustainable Mining were not included in the study. The website of the newly-titled journal (jsm.gig.eu) has an online archive of the contents of all the regular RRCMIME issues and one abstract issue. PPJs are available only in print form. The abstract issue was published in 2004 for a conference entitled, Conference Naturally Occurring Radioactive Materials IV. This issue includes 95 abstracts and 1 editorial. Full-texts of papers have been published independently from RRCMIME as proceedings.

RRCMIME includes 43 regular issues with 299 papers and 5 editorials with the number of papers each year being quite similar [Figure 1]. In 2006–2011, proceedings papers were also published and took the form of special issues. Eight-hundred six papers and 10 editorials were published in 29 special issues. The proceedings papers in RRCMIME make up 67.1% of the journal’s contents.
In RRCMIME, 1185 authors published their papers: Two hundred thirty-eight in regular issues, 808 in PPJ, and 219 in abstract issue. Sixty percent of all the papers were co-authored, mainly by 2 or 3 authors (589 papers). The addresses of the authors were not examined because the RRCMIME journal data are incomplete. Full-text papers were mainly published in Polish (85.83%). Each paper is provided with the title and abstract in English. Full-text papers in other languages make up 6.08% (English – 65 papers, Czech – 7 papers, and German – 1 paper).

Data Sources and Search Strategy

Our search was conducted from March to June 2014. Citation data came from four sources such as the WoS™ Core Collection, Scopus, BazTech, and GS. In the case of WoS, the following citation index was considered: Science Citation Index Expanded (2002–), Social Sciences Citation Index (2002–), and Conference Proceedings Citation Indexes (2002–). Search options and items used in all sources for this study are described in Table 1.

RCMIME is not indexed in WoS; therefore, we used the “Cited Reference Search” form to search for citations. Two strategies to search for citations could be used: (1) Determining the variations of the journal title and carrying out searches on that basis[8] or (2) searching by author name or article title. Given the multiword title of the RRCMIME journal, the editors using an English parallel title and the titles of conferences, we assumed that authors might have written the journal title in different ways, creating numerous variants.[13] The first strategy would involve determining prior variants of the journal title. As the habits of the authors citing RRCMIME were unknown (there is no research on this matter), we chose to search by author name.

In WoS, 1185 names of the authors of 1216 RRCMIME papers were used as search terms. In our opinion, this way of searching ensures greater accuracy when the journal title is missing or is an abbreviated form. We were able to verify some mistakes and get the citations by searching in different data sources and, in the case of co-authored papers, searching by the name of each author of the paper. The use of the names of all of the authors increased the number of searches but, in the case of an incorrect entry of the first author, it enabled the identification of the citation.

In Scopus, the search was conducted by author name using the “Author Search” feature and, after the results were displayed, the “view secondary documents” option was chosen. This strategy enabled maximum accuracy in finding papers published in journals that are not indexed in Scopus.

RRCMIME is a domestic journal and, that is, why it was also decided to search its citations in the BazTech database which indexes over 600 Polish technical-related journals. The contents of the database are also indexed in GS. Since 2006, each record entered into the database has been completed with references. Citations in BazTech are registered as plain text in the “Bibliografia” (references) field and, therefore, searching is carried out throughout the cited literature in a paper and covers all the elements of the bibliographic description.

The last data source we employed was GS. Searches were done directly in the web browser, and paper titles were used as search terms. If there were some doubts, authors’ names were also used. Citations in GS are displayed in two ways:
- As a link located under the cited article with a corresponding number of citations (“cited by”), which is a direct reference to the list of citing works;
- As a bold part of the publication, visible on the results list.

Our experiences are consistent with research by Meier and Conkling.[14] For the purposes of this analysis, all citations regardless of how they are interpreted by GS were taken into account. In light of the fact that citation
matching in GS is performed only on the basis of article titles, each result citation was additionally checked in the citing document.

Self-citations

Two aspects of self-citations were considered, author self-citations and journal self-citations. In the case of author self-citations, the definitions proposed by Snyder and Bonzi[15] and Aksnes[16] were used, according to which a self-citation occurs whenever the set of co-authors of the citing and cited article share at least one author. For journal self-citations, citations in RRCMIME and the Journal of Sustainable Mining were also counted.

International Diffusion

International diffusion of selected Polish journals was studied by Racki.[5] On the basis of the first author's country and its changes, he determined the international scope of the impact of the Geological Quarterly. A different method, which was used in our study, was proposed by Kortelainen.[17] According to this method, a citation is regarded as domestic when at least one author is Polish, regardless of the nationality of the other citing authors. When none of the authors are Polish, the citation is considered foreign. In the case of multi-authored papers, each country is given 0.5 points if the authors represent two countries, 0.33 points if they represent three countries, etc. As a data source for the analysis of citing authors' countries, we chose Scopus and WoS. Therefore, the analysis included only part of the citations. BazTech records created before the end of June 2013 only had the first author's address. The use of the first author's address does not show the full picture of scientific cooperation so we did not utilize this source. GS does not index such data; therefore, any citing source should be checked manually.

RESULTS AND DISCUSSION

Citation Count from Different Sources

From the four sources, 800 unique citations were received. The detailed bibliometric characteristics of RRCMIME obtained from all sources and with division on individual sources are presented in Table 2.

The highest amount of citations (631) was found in the Polish BazTech database. The second largest source was GS with 586 citations. The contents of BazTech are comprised in GS, so one might expect that the result would be higher than in the Polish database as GS covers the entire collection of BazTech plus data from other sources. This hypothesis, however, was not proven. In GS, 493 citations (84%) recorded in BazTech were found. The remaining 138 items could not be found in GS. This observation suggests a careful and critical approach to the results obtained by GS.

In Scopus, 165 citations were found and in WoS –156. A similarity of the results can be seen in the two sources. This is not coincidence since most journals citing RRCMIME are indexed in both databases. In this case, 137 items (almost 83%) were recorded in both databases.

Among 800 citations found, 6.5% (52 items) are present in all four sources.

The first citations appeared in the 3rd year of the existence of the journal, i.e. in 2004. In 2004–2006, the journal was cited 23 times. Significant growth can be observed in 2008 (113 citations), and the upward trend was maintained until 2013. Due to incomplete data, 2014 was deemed inconclusive.

Figure 2 shows the number of citations of RRCMIME in each source in the years 2004–2014.

In Figure 2, it can be observed that the distribution of citations for each source in the subsequent years is very similar. With regards to the number of citations, the best years were 2011–2013 when the journal received more than the half of all of its citations.

Of the 1216 papers, over 31% (380) have been cited at least once, of these 38 have five or more citations.

The analysis shows that the most cited papers were published in 2003 and from 2005 to 2007 (45–55% of the papers). The weakest year in this respect was 2004. The low figure (12%) was highly influenced by one issue with several dozen abstracts of conference proceedings which were cited only 3 times. The full-text papers of the conference, which were published as a monograph, were much more frequently cited. Figure 3 shows the ratio of cited papers to not cited papers.

The papers receiving the most citations were published in 2007. One-hundred sixty articles, which were published in four regular and five special issues, received 187 citations (over 23% of the total). The average citation rate per paper published in 2007 was 1.17 and 2005 had an even greater
average at 1.52. Twenty-nine articles were cited 44 times. The papers published in 2004 and 2012 had the lowest impact in literature. The average in these 2 years was 0.24 and 0.29, respectively.

PPJ received the most citations. Among the 10 most cited volumes, eight are PPJ. They obtained a third of all RRCMIME citations. Papers from 30 issues were cited a total of 308 times. Articles from PPJ and issue with abstracts have 61.5% of all citations, cited 492 times. However, it is worth mentioning that PPJ contained 3 times more papers than regular issues; hence, the average citation ratio per paper in a regular issue is greater and amounts to 1.01 while, in a PPJ, it is 0.54. The lower citation of PPJ confirms the research carried out by González-Albo and Bordons.[18]

The highest average number of citations per paper was issue number 1, 2003. Four articles from this volume were cited 19 times (an average of 4.75 times). The highest average for a special issue (number 1, 2008) was 1.67 per paper. Articles in special issue number 3, 2007 received the most citations (59). They were the results of a conference held in the same year. For all 22 issues, the average citation ratio per paper is more than one.

### Cited and Citing Authors

There were 1185 authors in RRCMIME. Four hundred fifty-three of them wrote two or more papers. Papers of 440 (37%) authors were cited. Articles of eight authors were cited more than 20 times. Józef Dubiński received the most citations; his papers had 39 citations including 14

---

**Table 2: Bibliometric characteristics of RRCMIME based on individual sources and unique citations received from all sources**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>WoS</th>
<th>Scopus</th>
<th>BazTech</th>
<th>GS</th>
<th>All sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of citations in total and in</td>
<td>156</td>
<td>165</td>
<td>631</td>
<td>586</td>
<td>800</td>
</tr>
<tr>
<td>Regular issues</td>
<td>67</td>
<td>75</td>
<td>233</td>
<td>231</td>
<td>308</td>
</tr>
<tr>
<td>PPJ</td>
<td>86</td>
<td>87</td>
<td>397</td>
<td>354</td>
<td>489</td>
</tr>
<tr>
<td>Abstract issue</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Average number of CPP</td>
<td>0.128</td>
<td>0.135</td>
<td>0.274</td>
<td>0.482</td>
<td>0.658</td>
</tr>
<tr>
<td>Number of citing papers</td>
<td>121</td>
<td>130</td>
<td>480</td>
<td>470</td>
<td>620</td>
</tr>
<tr>
<td>Number of citing journals (number of citations)</td>
<td>43 (141)</td>
<td>51 (150)</td>
<td>72 (631)</td>
<td>82 (635)</td>
<td>118 (738)</td>
</tr>
<tr>
<td>Number of other citing items (number of citations)</td>
<td>8 (15)</td>
<td>13 (15)</td>
<td>N/A</td>
<td>42 (51)</td>
<td>50 (62)</td>
</tr>
<tr>
<td>Percentage of papers with author self-citation</td>
<td>58.33</td>
<td>58.2</td>
<td>61.97</td>
<td>61.26</td>
<td>59.25</td>
</tr>
<tr>
<td>Percentage of journal self-citation</td>
<td>N/A</td>
<td>N/A</td>
<td>13.63</td>
<td>13.31</td>
<td>11.5</td>
</tr>
<tr>
<td>H-index</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>8</td>
</tr>
</tbody>
</table>

CPP=Citations per paper, NA=Not available, GS=Google Scholar

**Figure 2: Number of citations in each database and unique citations received from all sources**
self-citations. The next groups are as follows: 28 authors with 10–19 citations, 58 authors with 5–9 citations, 74 authors with two citations, and 188 (41%) with one citation. These results were based upon whole counting scheme which assigns contribution to authors and co-authors equally.

More than half of the publications (60%) were created as a result of cooperation between at least two authors. Three articles were written by teams of 10–13 people. To determine the proportional contribution of citation for each of the co-authors, the citations for each of them were counted (fractional counting scheme). The person who received the most citations per author was Grzegorz Mutke (21.85 citations). Only two other authors received more than 20 citations. Twelve authors were cited 11–18.5 times; 23 authors were cited 5–9.25 times; 30 authors were cited 3–4.5 times; 34 authors were cited 2–2.93 times. Three quarters of authors received the average of 0.1–1.83 citations including 225 authors (slightly more than a half of all cited authors) with the number of citations being <1. On the basis of the data, it can be concluded that regardless of the method of calculating citations, the top of the ranking is very similar.

Almost 60% of citing papers are author’s self-citations. In teams of authors, there was at least one person who cited their own publication.

Citing Journals, Other Items, and Self-citations

We divided citing sources into: International Journals Indexed in WoS, other non-Polish journals, Polish domestic journals, other types of publications, and journal self-citations [Table 3]. About 80.75% (646) of all citations came from 117 journals. Other types of publications made up 7.75% and journal self-citations – 11.5% (92 citations).

The highest number of citations comes from domestic journals (485 citations from 62 journals). The Polish local journals that cited RRCMIME the most times are Przegląd Górnicy (title in English: Mining Review) – 17.13%, Wiadomości Górnice (title in English: Mining News) – 8.25%, and 5 other Polish journals (a total of 19.5% of all citations). The remaining 55 Polish domestic journals provided 15.75% of citations.

International Journals Indexed in WoS (48 journals) accounted for 19% (152) of the total citation number. Nearly half of that number (45.39%, 69 citations) were provided by two journals with the editors based in Poland (Archives of Mining Sciences [AMS], and Gospodarka Surowcami Mineralnymi-Mineral Resources Management [GSM]) which makes up 8.23% of the total. Four journals cited RRCMIME 4–7 times each, 19 journals – 2 or 3 times, and the remaining 23 journals – once. These results prompted us to examine how far away RRCMIME is from two Polish Journals indexed in the WoS and who cite RRCMIME most often.

The number of citations and self-citations of authors of papers published in 2007–2012 were compared, AMS and GSM have been indexed in WoS since 2007, and RRCMIME ceased to be published in 2012. The most often cited journal (with regards to the average number of citations per paper) from the three compared was AMS (1.6 CPP), then GSM (0.87 CPP). The CPP ratio for RRCMIME was 0.12. In all three journals, a high degree of author self-citations can be seen: 68.58% (AMS), 59.63% (RRCMIME), and 55.76% (GSM). The results show why RRCMIME did not meet the requirements for journals indexed in WoS.

Citations from 7 other international journals (with the editors located in a country other than Poland) accounted...
for 1.13% (9 citations). These journals were indexed either in Scopus (3 journals) or in GS (4 journals).

Other types of publications provided 62 citations. These were mainly conference proceedings (3.75% of all citations). Monographs provided 1.75% of citations, Ph.D. theses –0.63%, and reports –0.25%. There were also 10 other publications whose types could not be determined. RRCMIME citations identified in this group of publications mostly came from GS. These citations were found in the vast majority (72.58% citations in this group) only in GS. It should be noted that RRCMIME citations in other types of publications are of marginal importance.

Citing Countries

We prepared information on the countries of the citing authors’ origin on the basis of data from Scopus and WoS. We included 184 citations from 145 citing articles. The vast majority of citations were domestic (92.26%) which is consistent with other studies on Polish-language journals not indexed in WoS or Scopus. The remaining 7.74% come from 7 countries, including the Czech Republic (7 citations); India (6 citations); Canada (4 citations); Romania (2 citations); and France, Norway, and Egypt (1 citation each).

The results confirm the domestic impact of RRCMIME because, even when RRCMIME is cited in international journals, the citations are mostly by authors from Poland. It stems from the fact that 86% of the articles published in RRCMIME were written in Polish. Another Polish journal (Geological Quarterly) that publishes full-texts in English received 5 times more citations from outside Poland which was 41% of the total number of citations.

Time from Publication to Citation

We determined how many years passed since the publication of the paper until it was cited in research literature. For 4 citations from GS, we were not able to find the year of citation. The longest amount of time after publication was after 11 years, which applies to two articles. The least amount of time, in which RRCMIME papers were cited 36 times (4.5% of the total citations), and without self-citations 12 times (3.7%), was in the same year of publication. RRCMIME papers received the highest number of citations in the 1st year after their publication (202 citations). This result was influenced significantly by author self-citations, which accounted for 75.74% of the citations. Excluding self-citations, RRCMIME papers received the most number of citations 2 years after publication (21%). Figure 4 presents the number of citations of RRCMIME counted in the time that elapsed between the publication year of a paper and the year it was cited including author with self-citations [Figure 4a] and without self-citations [Figure 4b].

The average time for all RRCMIME papers measured from the year of publication to citation is 3 years (2.96). If we separate conference proceedings from regular issues, then the average time for conference proceedings papers is 2.43 years and regular issue papers is 3.82 years.

CONCLUSIONS

RRCMIME as a Polish-language journal is cited by journals indexed in International and Interdisciplinary Databases (WoS and Scopus). One-third of the journals’ editors were in Poland. One in five citations of RRCMIME was found in one of these databases.

The ratio of RRCMIME citations in international journals: Local journals is 1:4, which indicates that RRCMIE has mainly domestic influence. The reason is the high number of full-text papers in Polish, which significantly narrows the journal’s audience. Another reason is the local nature of the topics covered in the articles. Some papers were
cited in foreign literature, but these citations were by Polish authors in more than 90% of cases.

The highest numbers of RRCMIME citations were contained in the BazTech database. BazTech proved to be the most complete source for bibliometric research on national journals in the field of technical sciences including mining sciences. Therefore, BazTech could act as the basis of the citation ranking of Polish Technical Journals.

The PPJ contained 67.1% and regular issues 25% of RRCMIME papers. The remaining 7.89% are abstracts. Although PPJ scored 61.5%, their citations per paper rate (CPP) was lower (CPP of PPJ = 0.54) than in the case of papers in regular issues (CPP of regular issue journals = 1.01).

The distribution of RRCMIME citations peaked in the 1st year of publication. These were in the vast majority author self-citations. Excluding self-citations, RRCMIME papers received the most citations 1 year after publication. This means that the authors citing RRCMIME articles used new publications.

An important observation was the fact that GS does not index citations that could be expected to be there. As BazTech contents are indexed in GS and the structure of the BazTech record does not change, one would expect full BazTech resources in GS. However, our research revealed a 16% drop in citations in GS, which could be found in BazTech. It is difficult to tell how many citations from other sources are ignored by GS. Taking this fact into account, we recommend national citation indexes for bibliometric analyses of domestic journals not indexed in WoS or Scopus. Without such sources, GS would be the best choice.

In 2013, the RRCMIME journal underwent a complete makeover. Not only was the journal title changed but also the language of publications and the editorial policy, that was changed from local to international. The changes are designed to enhance quality. Examples from literature show that after changing the title and the language of full-text papers the number of journal citations is increasing, and the ratio of self-citations to citations is decreasing. Our paper presents a case study of one journal in the field of mining sciences. The results can serve as a reference in future studies of the journal in a new form. The findings of future research may be an indication as to whether the adopted policy of internationalization of the journal is effective and increases the impact of the Journal in World Science.

Financial Support and Sponsorship

The research work was supported under the statutory activity of the Central Mining Institute in Katowice (Poland) within project number 10851434-072.

Conflicts of Interest

There are no conflicts of interest.

REFERENCES