



Article **Coal Modeling Investigations in International Collaboration in** the Light of Bibliometric Analysis of the Problem

Agnieszka Saramak and Daniel Saramak *D

Department of Environmental Engineering, AGH University of Science and Technology, 30-059 Kraków, Poland * Correspondence: dsaramak@agh.edu.pl; Tel.: +48-12-617-4559

Abstract: The article concerns an analysis of records registered in Web of Science (WoS) database related to the problem of coal modeling. All publications registered in the WoS from the time period 1951–2021 were analyzed, mostly in terms of international collaboration, merit content and research areas. It appeared that international scientific cooperation on this topic is differently considered, depending on the country. The leader in terms of the number of documents remains China, while the highest citation counts were gained by research teams, with the USA as the leader. Several sub-categories within research areas could be also divided on the basis of key words, while the most popular topic is connected with energetical aspects of coal utilization.

Keywords: coal; modeling; bibliometric analysis; mineral processing



Citation: Saramak, A.; Saramak, D. Coal Modeling Investigations in International Collaboration in the Light of Bibliometric Analysis of the Problem. Energies 2022, 15, 6040. https://doi.org/10.3390/en15166040

Academic Editor: Manoj Khandelwal

Received: 14 July 2022 Accepted: 18 August 2022 Published: 20 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

1. Introduction—Review of the Problem and the Scope

Practical utilization of coal is connected with many industrial branches. Excluding aggregates, coal is the raw material with the highest volume of mining production, reaching globally nearly 7 billion (10^9) tons annually (Figure 1). It is also a primary energy raw material with strategic meaning, despite global trends aiming at adopting more environmentally friendly methods of energy production.



Figure 1. Mining production of main types of coal in the period 1984–2020 [1].

It should also be pointed out that world coal consumption can be described as a steady but increasing trend, analyzing the time period of the recent four decades. This is also linked with a number of various research projects on coal and a number of scientific publications, as mentioned in the following section. Figure 1 presents the trend of the world production of coal over the last four decades.

New directions of investigations on coal are being carried out, including the development of new technologies and procedures that help in decreasing the emission of CO₂ and other gas and particulate matter pollutants. In fact, contemporary scientific investigations on coal are strictly related to the entire value chain of coal production, especially with issues of coal acquisition, problems of processing technology and aspects of its industrial utilization. Current trends in coal research also have their reflection in focusing on environmental aspects, including the above-mentioned problems of the emission of gas and dust pollutants, but not limited to the other pending problems, such as logistics and supply limitations [2–4]. Social and economic aspects are also present among the pressing problems concerning the topic [5]. In summary, the following general topics in coal research can be distinguished.

Mining activities, starting from geological investigations, deposit characterization and a variety of geotechnical problems, through to mine operation including a wide range of safety aspects, production effectiveness, process control and operations indirectly connected to the mine production system.

- Technology of upgrading coal, including preparatory operations and mechanical processing as well as coal beneficiation processes, together with new technologies aiming at the improvement of the qualitative properties of final products;
- Industrial and commercial utilization of coal, with low-emission combustion technologies, improving the energetic and economic effect of coal consumption;
- Environmental aspects that appear during the entire life cycle of coal;
- Economic issues.

It is worthwhile to note that the above topics are thematically intertwined and often cannot be analyzed separately, especially in more strategic investigations and projects. Therefore, there is always room for creating a new direction of research or more detailed development of the existing topics. Models related to coal research are in fact present in each of the above directions, but they differ significantly depending on their utilization area, methodological bases, mathematical apparatus used and the purpose.

Coal modeling for the purpose of this paper is understood as an approach that uses any kind of simulation, approximation, optimization, prediction and other evaluating procedures and tools correlated to coal in general. From a technical point of view, models very often use mathematical apparatus for the determination and description of a specific phenomenon, process or action, but a variety of descriptive models based on expert opinions and interactions, i.e., Delphi models and the AHP process [6], also exist. There are various classifications of general models and models dedicated to a specific topic, such as coal, regarding the area, time, technology and point in the coal value chain and many others. The authors' intention was not to limit the scope to specific types of models but to take into consideration all records connected with coal and modeling approaches.

Coal modeling in the stage of mining activities may concern very general and fundamental issues, such as mine operation, the estimation of mine servicing time, quality of raw material and others [7–9]. Geological investigations in underground coal mines show geotechnical models based on numerical analysis, a probabilistic approach and statistical analysis that characterize the stability of the rock mass [10–13]. Simulation models that control the flow of gas and particle mattes in underground mines are useful for improving the operation of ventilation systems and for contaminant removal. Developments of these models are possible through the utilization of computational fluid dynamic (CFD) tools, regression techniques and machine learning [14–16].

Coal processing technology is a field with many opportunities for the application of both simulation and optimization models. Simulation models based on the discrete elements method (DEM) and the mentioned CFD may bring a detailed description of the feed material behavior during the process in a device, which helps in the design of more effective machines [17–21], while the results of coal beneficiation in separation processes can be modeled by means of mathematical programming theory, regression models, neural networks and genetic algorithms [22–24].

Industrial utilization of coal mostly concerns investigations into obtaining an optimal amount of energy from the combustion of the raw material, together with the neutralization or at least limitation of the major drawbacks of coal fuel use, such as the emission of toxic gases, ashes and slags and increasing the greenhouse effect. Research on the development of more effective, environmentally and economic friendly technologies also has a reflection in contemporary science concerning the issue [25–31]. These investigative projects combine various disciplines, such as chemistry, physics, material sciences and engineering areas.

The thermal conversion of mixtures, i.e., coal–water (or coal–other medium), heat transfer, kinetics and other chemical or thermodynamic problems are covered [32–35]. Mathematical models are very popular within this topic, but numerical modeling, simulations and predictions are also among the utilized tools and methods [36–38]. Environmental aspects of coal modeling and the coal itself were always a significant

interest of researchers. The issue has gained even more importance as a result of higher environmental awareness in society, but also due to the ratification of various international agreements aiming at carbon emission reduction, greenhouse effect limitation and a general effort for the improvement of societies' quality of life. Coal extraction causes a number of hydrological consequences [39–42] and is a source of a significant amount of waste materials that interact with the environment, causing increased erosion, acid drainage and concentration of dissolved soils [43]. Models concerning the reduction of carbon emission in general and through a specific design of the mining process are also applied [44]. Mathematical modeling and simulations are widely used here, but also statistical and predictive models are also in use [45–48].

Economic models on coal concern the supply–demand balance, as well as supporting the decision-making in the rational planning of the development of an industry, more usually at a microeconomic scale [49,50]. Such models help in analysis through the development of general scenarios of sector growth, as well as forecasting the coal demands through the estimation of the required energy governance and consumption, considering different scenarios. Simulation and optimization models are developed, utilizing, among others, a theory of mathematical programming and auto-regression [51–54]

There are also publications presenting the results of a bibliometric analysis carried out in the field of coal, but they focus either on selective aspects, a limited time period or a narrower area, i.e., domestic industry [55–57], and that needs updating and expansion due to the increased number of publications issued in recent years (Figure 2). Bibliometric analyses are becoming popular methods of measuring impact for scientific publications on the specific topic [58,59]. Authors are aware, however, that the bibliometric analysis of investigations on coal appears a huge and challenging task. The multithreading of the topic that appeared during investigations made it impossible to analyze all significant findings in a manner concise enough to include it in a single paper. It also appeared that selected problems were so complex that it was not possible to find relevant conclusions without using more advanced tools than those built into a specific database. The problems concerned especially the analysis of cooperation and the merit content through authors' and editors' key words, as well as the cluster analysis. Therefore, minor simplifications to the issue were applied, and the authors decided to analyze publications registered only in one scientific database (the Web of Science database).

In summary, the main aim of analysis presented in the paper sis to address the following questions:

- What is the status of international collaboration within the topic?
- In what directions can research on coal modeling evolve?
- How are the publications clustered?
- What areas are dominant in investigations?



Figure 2. Number of publications in individual years.

Taking into account the above-mentioned issues, as well as the fact that the number of publications concerning coal grows year by year (Figure 2), it was assumed that the existing research within this area needed expansion and updating. Therefore, the general purpose of this comprehensive bibliometric analysis was to identify research patterns and trends in the scientific literature on coal modeling. The authors have tried to identify these patterns and trends through selected aspects, characterized in the following sections. Bibliometric analysis is a popular method of exploring and analyzing large amounts of scientific data that enables a comprehensive review of the literature in a selected field, identifying knowledge gaps, generating new research ideas and positioning the intended contribution of scientists to the field in question [60–62].

There is a general agreement between researchers that research quality cannot be characterized by a single element, especially when more detailed conclusions are expected to be drawn. Single metrics are of course helpful, but a particular indicator or index may fail to capture some significant and different aspects concerning the specific research or investigative program [63]. A variety of indicators can be used in the assessment of scientific research, and for the purposes of this paper, the following types of metrics were used:

- Metrics of productivity, including the number of research works in cooperation, number of publications, number of institutions involved, and others. These characteristics can be also refer to the time period or specific year;
- Metrics of impact, showing productivity metrics related to a specific author, institution, publication, country, year or others;
- Hybrid metrics, combining various measures, such as the number of publications/citations/ cooperations per affiliation/country/author. This group of measures captures both the productivity and impact of a specific item in a single number or figure. They can refer to the same or different research areas [64,65]. This group of metrics is also associated with the Hirsch index [66,67].

The above metrics can be also generated from different points of view, i.e., for individual (or a group of) journals, editors, institutions, countries, authors, types of documents and many others. A significant number of the obtained indices as a result in such a way may be confusing in the analysis, especially given that some metrics can be correlated between each other. In addition, a bibliometric analysis of key words or research areas can be confusing for the reason that too many searching outcomes can be obtained, sometimes with similar or even doubled meanings. Cluster analysis can be a good solution in such cases, because the results can be partitioned into clusters that contain words, phrases or metrics that are mutually strongly correlated. Clustering can be performed with the use of various methodologies and can be performed using of specific tools and software, such as CiteSpace. The LSI method (Latent Semantic Indexing) groups the outcomes (words or phrases) according to the main keyword, and results in each group are semantically relevant. The terms are also ranked using different algorithms, such as LLR (log-likelihood ratio) and MI (mutual information) [68].

Detailed characteristics of the methodology and procedures used in the data collection and analyses are characterized in Section 2.

2. Materials and Methods

The presented analysis was carried out according to the following steps.

- Step 1: Study of the problem (selection of the searching phrase, selection of scientific database, selection of searching options for the key word, selection of the time period);
- Step 2: Data collection (collection of data from the database, screening of the results, exporting the final list to external file in Excel format, export of the final list into dedicated software);
- Step 3: Bibliometric analysis (presentation of selected descriptive statistics, mapping the networks of cooperation, merit content analysis);
- Step 4: Discussion of results and remarks.

The selection criteria in step 1 were defined as a combination of two words: "coal" and "modeling". These two words were combined with the "AND" operator in the searching option in order to narrow down the searching area only to records related to coal and covering the aspects of modeling. The word "modeling", according to British English spelling, was selected. We used the "all field" option in searching for the term "coal" AND "modeling", which means the phrase was searched in the title, abstract and keywords. It was also decided that all types of publications would be used. The timespan was preliminarily set as "all years" (default from 1900 ongoing), but it appeared that the first publication registered in the database was published in 1951. The authors also wanted to finish the list with a full year; therefore, the timespan was finally set as 1951–2021.

As mentioned in Section 1, the Web of Science database was selected, and the core collection option was used. According to the authors' opinion, both WoS and SCOPUS are equal regarding the quality of presented papers, and most publications that are indexed in SCOPUS are also present in WoS. Preliminary analysis confirmed that only a minor percentage of obtained records was different in each database, and both sources can in fact be used to complete satisfactory searching results. Among the minor advantages of WoS could be the fact that an insignificantly higher percentage of publications registered in it has the granted IF index compared with the SCOPUS database. This factor is very popular for the quantification of top-quality publications. A major drawback of WoS, however, is connected with technical issues of data gathering. There are limitations in the maximum number of records that can be downloaded from the database in a single file; therefore, the export had to be executed in several files, which were then bound together. This caused some problems, especially when the searching data were exported to the external software CiteSpace for use in further analyses. Up to 1000 records could be exported at a time; therefore, the completion of the overall input file consumed a considerable amount of time.

These restrictions were established to protect the the WoS data from excessive and automatic data copying and downloading, and advanced analyses were carried out by utilizing the InCites tool. This professional bibliometric software (CiteSpace 6.1), created by SourceForge (San Diego, CA, USA), was used mainly for in-depth analyses of key words, cooperation and clusters. This tool provides professional evaluations of patterns and trends in scientific literature and is utilized worldwide by academic and non-academic researchers. The software helps to address questions such as the following: what are major research areas connected to data, what are main connections between these areas, which areas are more active and more productive, which are key papers (key institutions, authors, countries) related to the problem/area and others. This tool enables the researcher to detect and track the knowledge on a certain topic and visualize networks of topic words and titles, collaborations and clustering. The specific organization of data makes it possible to follow the changes in the structures of collaboration, which are under constant modification over

time. The work of Chen [69] was one of the initial publications concerning the problem, but many other examples of publications can be found as developed versions of the CiteSpace software were released [70–72].

The raw data for analysis were obtained from the Web of Science database. The Web of Science Core Collection database was selected, and the primary searching criterion was set as "COAL & MODELING". We chose the all-fields option in the searched documents, and the analyzed period included the years 1951–2021. The searching was performed on 24 February 2022. Table 1 summarizes the preliminary characteristics of the searching procedure.

Table 1. Summary of searching design.

Item	Description
Searching database	Web of Science
Type of database	WoS core collection
Searching words	Coal, modeling
Searching operator	AND
Searching fields	Title, abstract, key words
Language	All
Nationality (region)	All
Type of publications	All peer-reviewed papers indexed
Timespan	1951–2021
Software	CiteSpace 6.1

After the preliminary screening of the searching results, the final list was exported to an external Excel file. Separate files had to be exported manually each time for a selected criterion. The following criteria were used in the preparation of the searching results.

- Year,
- WoS category,
- Type of documents,
- Affiliation,
- Title (journal),
- Publisher,
- Research area,
- Funding agency,
- Country.

3. Results and Analysis

As many as 53,372 records in total were returned after searching the WoS database. The phrase "COAL" itself returned 230,946 records, while the term "MODELING" returned 10,773,552. The obtained records were analyzed according to cooperation, authors' and editors' key words as well as clustering. Overall characteristics are provided in Table 2.

Table 2. General characteristics of searching results.

Item	Description	
Number of documents	53,372	
Numbers of titles (journals)	8107	
Number of countries	148	
Number of affiliations	16,147	
Number of publishers	1424	
Number of WoS categories	224	
Number of research areas	136	
Number of funding agencies	23,169	
Number of types of documents	19	
Number of years	62	

The numbers of years do not match the used timespan (1951–2021), for the reason that no publications were registered in the following years: 1952–53, 1958, 1961–65 and 1967. Some records appeared incomplete, as selected details in some categories were not provided. Table 3 below summarizes identified cases.

Category	Number of Incomplete Records
Title (journal)	0
Country	411
Affiliation	450
Publisher	3
WoS category	40
Research area	40
Funding agency	23,356
Type of document	0
Year	0

Table 3. Incomplete records in individual categories.

Incomplete records in the category *funding agency* do not mean that a file is corrupted, but it rather indicates that a publication was not financed through a dedicated project or entity. Therefore, these records were not considered as corrupted. Special attention was paid to the WoS categories. Certain categories were connected with medicine, humanistic sciences and other different areas (i.e., philosophy, religion, psychology, language sciences). In total, 95 WoS categories, containing 597 records, were considered as not connected with the topic and were removed from further analysis. In the case of the *research area* category, 70 out of 136 were left for analysis, and 420 records were removed from the searching results. Considering the fact that some incomplete records were sometimes in common for more than one category and that there were also the same publications in the removed WoS categories and research areas, 1580 publications were removed from the searching results, constituting nearly 3% of the total.

3.1. Main Outcomes of Cooperation

The analysis of cooperation was provided from the scope of the two countries with the highest number of documents: China and the USA. Figure 3 shows the number of common publications for China (Figure 3a) and other countries on the list, while Figure 3b visualizes the same cooperation but for the USA. Both figures include the 15 top countries that cooperate the most with both countries. Descriptions related to a specific country or region were provided by the authors originally, and different characterizations (i.e., UK and England) result from this fact only and were not modified during analysis.



Figure 3. Number of documents as a result of cooperation for China (a) and the USA (b).

Chinese institutions cooperated most frequently with entities from the USA. This cooperation was effective in nearly 1600 documents. The second country, China, cooperated most frequently with Australia, reaching 1048 documents. International cooperation between China and the mentioned two countries covered over 50% of the total international cooperation of China. International cooperation from the USA was over 60 times lower, in terms of the number of publications, compared to China. Nearly 50% of cooperation of the USA was with China. It can also be noticed that both China and the USA cooperate roughly with the same countries, and their order, to some extent, reflects the sequence on the list of top publishing countries (Table 4).

Country	Percentage Share [%]
China	48.467
USA	15.799
Australia	6.766
England	3.931
Germany	3.89
India	3.672
Canada	3.298
Poland	3.129
Japan	2.728
Russia	2.599
France	1.81
Spain	1.802
Italy	1.609
Turkey	1.602
South Korea	1.486
Netherlands	1.165
Sweden	1.089
Iran	1.04
South Africa	1.021
Czech Republic	0.811

Table 4. Top 20 countries with the highest number of publications.

The analysis of cooperation according to the number of citations also shows that China reached almost 120,000 citations in total, as a result of cooperation with the top 15 countries. The USA reached a much lower number of citations, at nearly 20,000, which is approximately six times lower than for China. However, the comparison of these numbers with the total number of documents (Figure 3) shows that documents that are the result of bilateral cooperation between the USA and other countries resulted in a 10 times higher number of citations per document, on average, compared to China. Cooperation between China and the USA produced 32.5% of citations counted as international from the Chinese point of view. In the case of the USA, the analogical number reaches 40.8%. The United Kingdom, Australia, Canada and Germany are among the countries producing the top number of citations in international collaboration with China and USA. (Figure 4).

Interested information can be drawn when percentage of cited documents in international collaboration, are analysed (Figure 5). An average percent of cited documents from international collaboration between China and other countries equals 90.95%, wile for USA it reached 100%. The results show that all publications related to USA cooperation seem to be scientifically significant, as all of them gained at least one citation. But it is worth to mention that only nearly 10% of scientific publications, being the result of Chinese international collaboration, hasn't gained citations. On the bases of the above, an international cooperation of the two most significant countries is on the high scientific level and bring significant findings that can be applied in further investigations and practical activities.



Figure 4. International collaboration expressed in citation counts for China (**a**) and USA (**b**)—top 15 countries.



Figure 5. Percentage of cited documents in international collaboration for China (a) and USA (b).

Another insight into the problem may be achieved through the utilization of the Category Normalized Citation Index (CNCI). Normalization, in general, helps to analyze a single publication in a wider context. As an example, a paper's citation count does not provide the whole information, i.e., the publication period, type of document or the frequency of citation in a specific research area. The CNCI is calculated by dividing the number of citations by the expected citation rate for documents with the same document type, year of publication and subject area. When a document is assigned to more than one subject area, the harmonic average is used. Results are shown in Figure 6.

Further statistical characteristics of this index are presented in Table 5. Results show that the average value of the CNCI for the USA equals 13.65, which is over seven times greater than for China (CNCI_China = 1.91). This shows that in a larger context, the quality of citations for papers from the USA can be considered as more favorable than papers from China. The variability index, defined as the relation of the average to standard deviation (m/σ), is also lower for the USA (2.55), while this value for China equals 4.66. The USA reached the highest values for cooperation with Norway (23.05) and England (21.84), while the lowest values were for Australia (5.44) and India (6.63). Analogous results for China are as follows: South Korea (2.66) and Austria (2.63) have the highest values, Germany (1.41) and Canada (1.47) the lowest.



Figure 6. Category Normalized Citation Index (CNCI) for China (a) and USA (b).

Index	Symbol	China	USA
Average	m	1.91	13.65
Standard deviation	σ	0.41	5.35
Variability	V	4.66	2.55
First Quartile	Q1	1.57	8.71
Median	Me	1.7	13.69
Third Quartile	Q3	2.17	19.15

Fable 5. Statistical	characteristics	of CNCI indices	for China and	USA.

An analysis of cooperation was also performed from the scope of international publications, especially those with the highest number of citations. The below analysis was carried out on the basis of the 100 top cited papers issued in cooperation with more than one country (Table 6). The table shows a list of countries that have the highest number of publications as a result of international cooperation, and such publications are on the list of the top 100 most cited papers.

Table 6. Top five countries with highest cooperation.

Country	Number of Publications in Top 100 Cited	Number of Citations in Top 100 Cited
China	27	18,040
USA	50	28,538
Australia	18	11,598
England	12	8759
Germany	13	5593

Figure 7 shows a visualization of bilateral cooperation in the form of lines. The width of each line reflects the number of produced citations within a specific cooperation. Numbers of total citations are provided next to each shape, with the number of common publications in brackets. It can be seen that the USA predominates in terms of mutual bilateral cooperation. The country has both the highest number of publications in the top 100 most cited and the highest number of total citations.

In the case of trilateral cooperation (Figure 8), the USA also plays a leading role. Only three such publications according to the provided country affiliation were on the list of the top 100 most cited, and in each of the groups, scientists from the USA were involved. Similarly, as in the case of Figure 3, the width of the bordering line of a set denotes the number of citations gained for the group.



Figure 7. Role of individual countries in bilateral scientific cooperation publications.



Figure 8. Role of individual countries in trilateral scientific cooperation publications.

3.2. Key Words Analysis

An analysis of key words was carried out separately for authors' key words and for editorial key words. As many as 968 authors' key words, with a total number of 37,545 appearances, were distinguished. However, only 62 words appeared more frequently than or equal to 100 times. For editor's key words, in turn, 352 words or phrases were distinguished with a frequency greater than or equal to 100 times. As many as 2152 words were distinguished in total in both categories, with a total indications count of 185,454. A summary of both types of words is presented in Table 7, while visualizations are shown in Figures 9 and 10.

Authors' Key Words		Editor Key Words	
Specific Key Word	Percentage Share	Specific Key Word	Percentage Share
Numerical simulation	2.66%	Model	3.62%
CO_2 capture	0.94%	Coal	2.16%
Coal mine	0.92%	Simulation	1.22%
Coalbed methane	0.89%	Behavior	1.19%
Coal combustion	0.87%	Combustion	1.14%
Mathematical model	0.86%	Kinetics	0.95%
Source apportionment	0.84%	System	0.95%
Coal mining	0.82%	Flow	0.90%
Fly ash	0.77%	Performance	0.88%
Fluidized bed	0.73%	Pyrolysis	0.76%

Table 7. Most frequent author and editor key words.

The analysis of authors' key words shows that the most frequent phrase was "numerical simulation", with a 2.66% share. The latter words reached a much lower share and did not exceed 1%. For editor key words, in turn, two words were predominant: "model" with 3.62% and "coal" with 2.16%. The three further key words exceeded a 1% share.

The visualization of both types of key words indicates the combinations and networks in which individual key words were present. The layout also organizes these key words and shows the most frequent links among them. Analyzing Figures 9 and 10, it can be noticed that individual words have a tendency to group themselves in natural sets. To investigate this phenomenon in more detail, a clustering analysis was performed.



Figure 9. Visualization of authors' key words.

bioma pyrol emoval technolog water emission on dic ation energy nodel flow designation ethpressure behavior svstem permeability impact evolution CiteSpace prediction

Figure 10. Visualization of editors' key words.

Clustering analysis was used for the evaluation of author and editor key words as well as research areas. On the basis of authors' key words, the following main clusters were defined using the LSI labeling method, numbered from #0 to #10:

- #0: climate change; CO₂ emissions; energy efficiency; greenhouse gases; carbon caps renewable energy; energy policy; energy storage; electricity grid; Nansihu lake basin;
- #1: coal gasification; gasification kinetics; drop tube furnace; initial crude; rheological curve | activation energy; Friedmann model; refuse plastic fuel; Arrhenius factor; toc removal;
- #2: neural network; genetic algorithm; seismic tomography; evolutionary programming; simulated evolution | artificial neural network; predictive model; nonlinear regression model; heating value; cosine amplitude method;
- #3: longwall mining; gob-side entry; backfilling body; self-stabilization structure; roof-controlled backfill ratio | numerical modeling; discrete element method; selfstabilization structure; roof-controlled backfill ratio; cutting test;
- #4: coalbed methane; stress sensitivity; reservoir volume; multi-wing fractured well; multiple transportation mechanisms | hydraulic fracturing; gas seepage; stimulating reservoir volumes; triaxial testing; saline aquifer;
- #5: numerical simulation; simple gray gas model; gray gases model; weighted sum; filtration rate | coal combustion; emissions reduction; sulfur dioxide; health risk; coalfired power generation;
- #6: CO₂ capture; chemical looping combustion; coal power plants; heat exchanger; phase-change material | carbon capture; power plant; rapid temperature swing adsorption; hollow fiber; sensitivity analysis;
- #7: vitrinite reflectance; burial history; basin modeling; temperature history; saxony basin | proximate analysis; ultimate analysis; carbon nanotube; fractal analysis; sensitivity analysis;
- #8: source apportionment; factor analysis; chemical mass balance model; Lake Michigan; Milwaukee harbor | positive matrix factorization; indoor air quality; health risk assessment; sensitivity analysis; heavy metals;

- #9: coal mining; water environment; mine water recycling; bi-level multi-objective programming; fractured coal | sensitivity analysis; chemical reaction kinetics; pollution reduction; percolation theory; potential index
- #10: sustainable development; circular economy; coal industry; coal mining; critical factors | coal enterprises; supply chain management; functional framework; procedure accomplishment; c-erp system.

Selected results of the clustering analysis are presented in Figure 11. Each cluster in the figure is marked with a different color and horizontal axes corresponding to the time period. The size of individual nodes on each axis denotes the concentration of phrases accounted to an individual cluster, while colored arched lines show items between clusters. It can be noticed that clusters from #0 to #5 are characterized with a relatively greater number of inter-cluster records. The highest concentrations of individual arches over horizontal axes are located in the same time periods for all clusters. However, this generally applies to the period 1998–2005, on average. Comparing this time period with Figure 2, it can be seen that, to some extent, this corresponds to two significant increases in the number of publications: the first following the year 1991 and the second starting from 2006. Such a rapid increase would be effective in a potentially higher number of new topics and investigative aspects that should be formulated or named with new words or phrases. Before this time, all these new descriptions had been sorted out, and some of them were classified as inter-clustering. Together with the development of a more concise description of clusters and their evolution, the numbers of inter-clustering records have been steadily reducing, and in recent years, only an insignificant number of them have been classified as inter-cluster.



Figure 11. Results of clustering analysis over time.

4. Discussion

Scientific papers registered in Web of Science are automatically assigned to one or more categories. All publications within the analyzed topic were assigned to over 200 categories in total, but approximately two-thirds of them have less than 100 counts. Figure 12 presents WoS categories that are connected with more than 100 publications. It is worthwhile to note that a single publication can be assigned to more than one WoS category; therefore, the total number of counts for all categories is higher than the number of analyzed publications. In example, only the top 10 most frequent WoS categories have a total share greater than 100% (Table 8).



Figure 12. Most frequent WoS categories.

Table 8. Top 10 most frequent Web of Science categories.

Web of Science Category	Percentage Share [%]
Energy Fuels	31.056
Engineering Chemical	23.788
Environmental Sciences	13.812
Geosciences Multidisciplinary	7.547
Mining Mineral Processing	7.163
Engineering Environmental	7.120
Thermodynamics	7.064
Engineering Mechanical	6.258
Materials Science Multidisciplinary	5.415
Chemistry Physical	3.948

Despite the huge number of WoS categories obtained for the analyzed topic, it is possible to group the categories into more general types. As a result, the authors propose to distinguish the following main areas related to the topic.

- Industrial utilization of coal and carbon products, mostly connected with energy production;
- Chemistry and chemical issues of coal treatment and industrial utilization;
- Environmental aspects affecting coal mining, processing and utilization;
- Technology of coal production, mostly on mining and mineral processing stage;

• Other engineering sciences related to coal different from mining and mineral processing. An analysis of the share of the above area shows that the area related to energy

has 43.5% of the share in total publications. The next area, chemistry of processes, is



represented by 32.7%—only 1% higher than environmental aspects (31.7%). Technology of coal processing covers nearly 24%, while the other engineering sciences cover 17%. A visualization of the joint effect of key words and categories is presented in Figure 13.

Figure 13. Authors' key words for WoS categories.

5. Conclusions

The analysis presented in this paper covered only selected issues related to the very wide topic of coal investigations. The authors focused mostly on international scientific collaboration and merit content through key words. The results brought the following conclusions.

- Many publications registered in the Web of Science database are the result of international cooperation, and the outcomes bring a significant input to the knowledge development within the issue. It can be also observed that different publications gain different citation counts, which may indicate that their interest may not be equal. However, the publications that were the results of bilateral or trilateral collaboration also gain a very high citation impact, which only confirms their top quality, as they are registered in a prestigious scientific database.
- Key words analysis showed that the most frequent research area is connected with the building of various simulation models and the utilization of mathematical tools in coal investigations. Environmental aspects and energy issues were also covered frequently in research. It is also interesting that authors' and editors' key words are to some extent convergent.
- Research area analysis content allows the distinguishing of several major sub-topics present in research: energetic aspects, coal chemistry, environmental issues, technology of processing and other engineering sciences, with the most frequent one related to energy (43.5%). Chemical and environmental aspects of coal modeling have approximately a 10% lower share.

• Clustering of results allowed us to distinguish the main groups according to merit content, and it appeared that a significant number of documents can be also classified as inter-cluster results.

On the basis of the obtained results, the following major directions in coal modeling can be characterized.

- Increasing the impact on environmental issues, especially in the development of various predictive models. Environmental aspects are gaining importance, also due to new legal regulations, often influencing on the macroscale, i.e., the limitation of CO₂ emissions, but the results of new investigative programs may contribute to the implementation of policies for the more effective protection of the environment, also on the global scale;
- Development of new technologies of coal preparation, targeting an increase in the amount of energy obtained from the mass unit of coal;
- More accurate prediction of the demand-supply balance, potential energy production and cost, both on a domestic and a more global scale.

Predictive models seem to be of significant importance for the coal industry, as their informative and descriptive aspects relating to both the short and long-term future may bring considerable benefits. The accuracy and precision of such models increases as many new efficient techniques are utilized, i.e., neutral networks. On the other hand, mathematical models can also be very precise, as specific technological processes that are the subject of modeling can be characterized in much more detail due to the knowledge development and better understanding of their course. There are still however gaps and uncertainties, and a number of aspects wait for further analysis and might be the subject of future investigations.

Author Contributions: Conceptualization, A.S.; methodology, A.S. and D.S.; software, D.S.; validation, A.S.; formal analysis, A.S. and D.S.; investigation, D.S. and A.S.; data curation, A.S.; writing original draft preparation, A.S.; writing—review and editing, D.S.; visualization, A.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Raw data can be found in the Web of Science database. Specific searching criteria apply (see Section 2).

Acknowledgments: Authors would like to thank the team of Bibliometric Analysis Group from the Main library of AGH University of Science and Technology for support in data collection.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. World Mining Data 2021, volume 36. Available online: https://www.world-mining-data.info/?World_Mining_Data_PDF-Files (accessed on 9 May 2021).
- Below, G.; Boland, N.; Savelsbergh, M.; Stuckey, P.J. Logistics optimization for a coal supply chain. J. Heuristics 2020, 26, 269–300. [CrossRef]
- Wang, H.; Xie, F.; Wang, L.; Li, J. Coal Logistics Demand Forecasting Based on Grey System in Shanxi Province. In Proceedings of the 2021 International Conference on E-Commerce and E-Management (ICECEM), Dalian, China, 24–26 September 2021; pp. 61–64. [CrossRef]
- 4. Zhang, Q.; Shen, H.; Huo, Y. An Evaluation Model of Green Coal Supplier for Thermal Power Supply Chain Based on PCA-SVM. *Math. Probl. Eng.* **2021**, 2021, 8827273. [CrossRef]
- Burchart-Korol, D.; Fugiel, A.; Czaplicka-Kolarz, K.; Turek, M. Model of environmental life cycle assessment for coal mining operations. *Sci. Total Environ.* 2016, 562, 61–72. [CrossRef]
- 6. Tumidajski, T.; Gawenda, T.; Niedoba, T.; Saramak, D. Directions of hard coal processing technological changes in Poland. *Miner. Resour. Manag.* **2008**, *24*, 245–258.

- Bernhardt, E.S.; Lutz, B.D.; King, R.S.; Fay, J.P.; Carter, C.E.; Helton, A.M.; Campagna, D.; Amos, J. How many mountains can we mine? Assessing the regional degradation of central Appalachian rivers by surface coal mining. *Environ. Sci. Technol.* 2012, 46, 8115–8122. [CrossRef]
- 8. Korski, J.; Tobór–Osadnik, K.; Wyganowska, M. Assessment of competitive ability of the Polish mining industry in the light of historical data. *Przegląd Górniczy* 2015, *6*, 9–15.
- 9. Korski, J.; Tobór–Osadnik, K.; Wyganowska, M. Reasons of problems of the polish hard coal mining in connection with restructuring changes in the period 1988–2014. *Resour. Policy* **2016**, *48*, 25–31. [CrossRef]
- Małkowski, P.; Ostrowski, Ł. Convergence Monitoring as a Basis for Numerical Analysis of Changes of Rock-Mass Quality and Hoek-Brown Failure Criterion Parameters Due to Longwall Excavation. *Arch. Min. Sci.* 2019, *64*, 93–118. Available online: http://archiwum.img-pan.krakow.pl/index.php/AMS/article/view/1166/872 (accessed on 3 August 2022).
- 11. Babets, D.; Sdvyzhkova, O.; Shashenko, O.; Kravchenko, K.; Cabana, E. Implementation of probabilistic approach to rock mass strength estimation while excavating through fault zones. *Min. Miner. Depos.* **2019**, *13*, 72–83. [CrossRef]
- Jing, L. A review of techniques, advances and outstanding issues in numerical modelling for rock mechanics and rock engineering. *Int. J. Rock Mech. Min. Sci.* 2013, 40, 283–353. [CrossRef]
- 13. Mayer, J.M.; Allen, D.M.; Gibson, H.D.; MacKie, D.C. Application of statistical approaches to analyze geological, geotechnical and hydrogeological data at a fractured-rock mine site in Northern Canada. *Hydrogeol. J.* **2014**, *22*, 1707–1723. [CrossRef]
- 14. Sasmito, A.P.; Birgersson, E.; Hung, C.; Arun, L.; Mujumdar, S. Some approaches to improve ventilation system in underground coal mines environment—A computational fluid dynamic study. *Tunn. Undergr. Space Technol.* **2013**, *34*, 82–95. [CrossRef]
- 15. Palei, S.K.; Das, S.K. Logistic regression model for prediction of roof fall risks in bord and pillar workings in coal mines. *Approach Saf. Sci.* **2009**, *47*, 88–96. [CrossRef]
- 16. Sarkarab, S.; Maitib, J. Machine learning in occupational accident analysis: A review using science mapping approach with citation network analysis. *Saf. Sci.* **2020**, *131*, 104900. [CrossRef]
- 17. Guerrero-Pérez, J.S.; Barraza-Burgos, J.M. A new mathematical model for coal flotation kinetics. Dyna 2017, 84, 143–149. [CrossRef]
- 18. Surowiak, A.; Brożek, M. A physical model of separation process by means of JIGS. *Physicochem. Probl. Miner. Process.* **2016**, *52*, 228–243.
- 19. Shahbazi, B.; Chelgani, S.C.H. Modeling of fine coal flotation separation based on particle characteristics and hydrodynamic conditions. *Int. J. Coal Sci. Technol.* **2016**, *3*, 429–439. [CrossRef]
- 20. Sand, A.; Rosenkranz, J.; Kuyumcu, H. DEM Modelling and Simulation of Coal Compaction by Stamping. In Proceedings of the Conference in Minerals Engineering, Luleå, Sweden, 2 February 2012.
- 21. Xie, Y.S.; Zhao, Y.S. Numerical simulation of the top coal caving process using the discrete element method. *Int. J. Rock Mech. Min. Sci.* 2009, *46*, 983–991. [CrossRef]
- 22. Öney, Ö. The Increase Of The Performance Of Ultrafine Coal Flotation By Using Emulsified Kerosene And The Prediction Of The Flotation Parameters By Random Forest And Genetic Algorithm. *Arch. Min. Sci.* **2019**, *64*, 119–130.
- 23. Jamróz, D. Application of multi-parameter data visualization by means of autoassociative neural networks to evaluate classification possibilities of various coal types. *Physicochem. Probl. Miner. Process.* **2014**, *50*, 719–734.
- 24. Golshani, T.; Jorjani, E.; Chelgani, S.C.; Shafaei, S.Z.; Heidari Nafechi, Y. Modeling and process optimization for microbial desulphurization of coal by using a two-level full factorial design. *Int. J. Min. Sci. Technol.* **2013**, *23*, 261–265. [CrossRef]
- 25. Hassea, C.H.; Debiagia, P.; Wena, X.; Hildebrandt, K.; Vascellarib, M.; Faravelli, T. Advanced modeling approaches for CFD simulations of coal combustion and gasification. *Prog. Energy Combust. Sci.* **2021**, *86*, 100938. [CrossRef]
- 26. Salomatov, V.; Kuznetsov, G.; Syrodoy, S.; Gutareva, N. Mathematical and physical modeling of the coal–water fuel particle ignition with a liquid film on the surface. *Energy Rep.* **2020**, *6*, 628–643. [CrossRef]
- Jamróz, D.; Niedoba, T.; Surowiak, A.; Tumidajski, T.; Szostek, R.; Gajer, M. Application of multi-parameter data visualization by means of multidimensional scaling to evaluate possibility of coal gasification. *Arch. Min. Sci.* 2017, 62, 445–457. [CrossRef]
- Strizhak, P.; Vershinina, A.; Yu, K. Maximum combustion temperature for coal-water slurry containing petrochemicals. *Energy* 2017, 120, 34–46. [CrossRef]
- 29. Thomas, D.J. Finding a future for clean coal and CO2 storage technology. Fuel 2017, 195, 314–315. [CrossRef]
- Marciniak-Kowalska, J.; Niedoba, T.; Surowiak, A.; Tumidajski, T. Multi-criteria evaluation of coal properties in terms of gasification. Arch. Min. Sci. 2014, 59, 677–690. [CrossRef]
- 31. Longwell, J.P.; Rubin, E.S.; Wilson, J. Coal: Energy for the future. Prog. Energy Combust. Sci. 1995, 21, 269–360. [CrossRef]
- 32. Rahman, M.; Pudasainee, D.; Gupta, R. Review on chemical upgrading of coal: Production, processes, potential applications and recent developments. *Fuel Process. Technol.* **2017**, 158, 35–56. [CrossRef]
- Kuznetsov, G.V.; Salomatov, V.V.; Syrodoy, S.V. Numerical simulation of ignition of particles of a coal–water fuel. *Combust. Explos.* Shock Waves 2015, 51, 409–415. [CrossRef]
- 34. Messerle, V.E.; Karpenko, E.I.; Ustimenko, A.B. Plasma assisted power coal combustion in the furnace of utility boiler: Numerical modeling and full-scale test. *Fuel* **2014**, *126*, 294–300. [CrossRef]
- Koornneef, J.; Junginger, M.; Faaij, A. Development of fluidized bed combustion—An overview of trends, performance and cost. Prog. Energy Combust. Sci. 2007, 33, 19–55. [CrossRef]
- Alekseenko, S.V.; Kuznetsov, V.A.; Mal'tsev, L.I.; Dekterev, A.A.; Chernetskii, M.Y. Analysis of combustion of coal-water fuel in low-power hot-water boiler via numerical modeling and experiments. J. Eng. Thermophys. 2019, 28, 177–186. [CrossRef]

- Orlova, K.Y.; Lebedev, B.V. Research of power fuel low-temperature vortex combustion in industrial boiler based on numerical modelling. *MATEC Web Conf.* 2017, 92, 01003. [CrossRef]
- Wen, X.; Luo, K.; Luo, Y.; Kassem, H.I.; Hanhui, J.; Fan, J. Large eddy simulation of a semi-industrial scale coal furnace using non-adiabatic three-stream flamelet/progress variable model. *Appl. Energy* 2016, 183, 1086–1097. [CrossRef]
- 39. Viney, N.R.; Post, D.A.; Crosbie, R.S.; Peeters, L. Modelling the impacts of future coal mining and coal seam gas extraction on river flows: A methodological framework. *J. Hydrol.* **2021**, *596*, 126–144. [CrossRef]
- 40. Lechner, A.M.; McIntyre, N.; Witt, K.; Raymond, C.M.; Arnold, S.; Scott, M.; Rifkin, W. Challenges of integrated modelling in mining regions to address social environmental and economic impacts. *Environ. Modell. Softw.* **2017**, *93*, 268–281. [CrossRef]
- 41. Guo, Q.; Su, N.; Yang, Y.; Li, J.; Wang, X. Using hydrological simulation to identify contribution of coal mining to runoff change in the Kuye River Basin, China. *Water Resour.* **2017**, *44*, 586–594. [CrossRef]
- 42. McIntyre, N.; Ballard, C.; Bruen, M.; Bulygina, N.; Buytaert, W.; Cluckie, I.; Dunn, S.; Ehret, U.; Ewen, J.; Gelfan, A.; et al. Modelling the hydrological impacts of rural land use change. *Hydrol. Res.* **2014**, *45*, 737–754. [CrossRef]
- 43. Bian, Z.; Inyang, H.I.; Daniels, J.L.; Otto, F.; Struthers, S. Environmental issues from coal mining and their solution. *Min. Sci. Technol.* 2010, 20, 215–223. [CrossRef]
- 44. Li, B.; Shi, Y.; Hao, J.; Ma, C.; Pang, C.; Yang, H. Research on a Carbon Emission Calculation Model and Method for an Underground Fully Mechanized Mining Process. *Energies* **2022**, *15*, 2871. [CrossRef]
- Post, D.A.; Crosbie, R.S.; Viney, N.R.; Peeters, L.J.M.; Zhang, Y.; Herron, N.; Janardhanan, S.; Wilkins, A.; Karim, F.; Aryal, S.; et al. Impacts of coal resource development in eastern Australia on groundwater and surface water. *J. Hydrol.* 2020, 591, 125281. [CrossRef]
- Sreekanth, J.; Crosbie, R.; Pickett, T.; Cui, T.; Peeters, L.; Slatter, E.; Northey, J.; Merrin, L.E.; Davies, P.; Miotlinski, K.; et al. Regional-scalemodelling and predictive uncertainty analysis of cumulative groundwater impacts from coal seam gas and coal mining developments. *Hydrogeol. J.* 2020, 28, 193–218. [CrossRef]
- Peeters, L.J.M.; Pagendam, D.E.; Crosbie, R.S.; Rachakonda, P.K.; Dawes, W.R.; Gao, L.; Marvanek, S.P.; Zhang, Y.; McVicar, T.R. Determining the initial spatial extent of an environmental impact assessment with a probabilistic screening methodology. *Environ. Mod. Softw.* 2018, 109, 353–367. [CrossRef]
- Viney, N.R.; Vaze, J.; Wang, B.; Zhang, Y.; Yang, A.; Vleeshouwer, J.; Ramchurn, A.; Frost, A. Intercomparison of methods for regionalising rainfall-runoff model predictions over large domains. In Proceedings of the Hydrology and Water Resources Symposium, Perth, Australia, 24–27 February 2014; pp. 970–977.
- Takaishvili, L. Simulation models to study development of the coal industry as energy sector component. *E3S Web Conf.* 2019, 77, 02004. [CrossRef]
- 50. Suwała, W. Modelling adaptation of the coal industry to sustainability conditions. Energy 2008, 33, 1015–1026. [CrossRef]
- Tomasi, S.; Gantioler, S. Innovative Approaches to Energy Governance: Preliminary Quantitative Insights from the Literature. In Smart and Sustainable Planning for Cities and Regions. SSPCR 2019. Green Energy and Technology; Springer: Cham, Switzerland, 2021; pp. 277–290.
- 52. Manowska, A. Use of autoregressive models to estimate a demand for hard coal. *Environ. Econ.* **2018**, *18*, 975–982.
- Agafonov, G.V.; Sokolov, A.D.; Takaishvili, L.N. Modeling of coal industry development. *Proc. RAS Power Eng. J.* 2011, *6*, 159–165.
 Kulshreshtha, M.; Parikh, J.K. Modeling demand for coal in India: Vector autoregressive models with cointegrated variables. *Energy* 2000, 25, 149–168. [CrossRef]
- 55. Zhang, K.; Wang, Q.; Liang, Q.M.; Chen, H. A bibliometric analysis of research on carbon tax from 1989 to 2014. *Renew. Sustain.* Energy Rev. 2016, 58, 297–310. [CrossRef]
- 56. Chen, D.; Liu, Z.; Zhao, H.L.; Webber, M.; Chen, J. Bibliometric and visualized analysis of emergy research. *Ecol. Eng.* **2016**, *90*, 285–293. [CrossRef]
- 57. Peng, B.; Guo, D.; Qiao, H.; Yang, Q.; Zhang, B.; Hayat, T.; Alsaedi, A.; Ahmad, B. Bibliometric and visualized analysis of China's coal research 2000–2015. J. Clean. Prod. 2018, 197, 1177–1189. [CrossRef]
- 58. Boudry, C.; Baudouin, C.; Mouriaux, F. International publication trends in dry eye disease research: A bibliometric analysis. *Ocul. Surf.* **2018**, *16*, 173–179. [CrossRef]
- 59. Li, J. Citation analysis: Comparison of Web of Science, Scopus, Scifinder, And Google Scholar. J. Electr. Res. Med. Libr. 2010, 7, 196–217. [CrossRef]
- 60. Donthu, N.; Kumar, S.; Mukherjee, D.; Pandey, N.; Lim, W.-M. How to conduct a bibliometric analysis: An overview and guidelines. *J. Bus. Res.* 2021, 133, 285–296. [CrossRef]
- 61. Janik, A.; Ryszko, A.; Szafraniec, M. Exploring the Social Innovation Research Field Based on a Comprehensive Bibliometric Analysis. J. Open Innov. Technol. Mark. Complex. 2021, 7, 226. [CrossRef]
- 62. Janik, A.; Ryszko, A.; Szafraniec, M. Scientific Landscape of Smart and Sustainable Cities Literature: A Bibliometric Analysis. *Sustainability* **2020**, *12*, 779. [CrossRef]
- 63. Franceschet, M. A cluster analysis of scholar and journal bibliometric indicators. J. Am. Soc. Inform. Sci. Technol. 2009, 60, 1950–1964. [CrossRef]
- 64. Egghe, L. Theory and practice of the g-index. Scientometrics 2006, 69, 131–152. [CrossRef]
- 65. Batista, P.D.; Campiteli, M.G.; Konouchi, O. Is it possible to compare researchers with different scientific interests? *Scientometrics* **2006**, *68*, 179–189. [CrossRef]

- 66. Hirsch, J.E. An index to quantify an individual's scientific research output. *Proc. Natl. Acad. Sci. USA* 2005, 102, 16569–16572. [CrossRef] [PubMed]
- 67. Katsaros, C.; Manolopoulos, Y.; Sidiropoulos, A. Generalized h-index for disclosing latent facts in citation networks. *Scientometrics* **2006**, *72*, 253–280.
- 68. Salton, G.; Wong, A.; Yang, C.S. A Vector Space Model for Automatic Indexing. Commun. ACM 1975, 18, 613–620. [CrossRef]
- 69. Chen, C. Searching for intellectual turning points: Progressive knowledge domain visualization. *Proc. Natl. Acad. Sci. USA* 2004, 101, 5303–5310. [CrossRef]
- 70. Chen, C.M. CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *J. Am. Soc. Inform. Sci. Technol.* **2006**, *57*, 359–377. [CrossRef]
- 71. Chen, C.; Ibekwe-SanJuan, F.; Hou, J. The structure and dynamics of co-citation clusters: A multiple-perspective co-citation analysis. J. Am. Soc. Inform. Sci. Technol. 2010, 61, 1386–1409. [CrossRef]
- 72. Chen, C.; Hu, Z.; Liu, S.; Tseng, H. Emerging trends in regenerative medicine: A scientometric analysis in CiteSpace. *Expert Opin. Biol. Ther.* **2012**, *12*, 593–608. [CrossRef]