TÜBİTAK

Turkish Journal of Veterinary and Animal Sciences

http://journals.tubitak.gov.tr/veterinary/

Research Article

Turk J Vet Anim Sci
(2021) 45: 833-840
© TÜBİTAK
doi:10.3906/vet-2001-103

Trend topics in animal science: a bibliometric analysis using CiteSpace

Fatma YARDİBİ*, Mehmet Ziya FIRAT, Emine ÇETİN TEKE
Department of Animal Science, Faculty of Science, Akdeniz University, Antalya, Turkey

Received: 30.01.2020 • **Accepted/Published Online:** 18.08.2021 • **Final Version:** 26.10.2021

Abstract: The aim of this study was the identification of trends topics in animal science in the last five years using bibliometric analysis. The research data consisted of 6972 studies published between 2015 and 2019 in the top five journals of animal science field, according to the Journal Citation Reports. The journals were analyzed in terms of number and types of publications, author, institution, country productivity, citation analysis, and citation burst. In the study, emerging trends and animal science intellectual structures were visualized with social network analysis. The evidence revealed in this study suggests that 'genomic prediction' is the most effective field of study in animal science field. 'Growth performance', 'Staphylococcus aureus', and 'Genomic prediction' were found as active clusters, and these topics may become popular in the future. Moreover, as a result of the word analysis conducted on the works made in the field, it was found that most repeated words are dairy cow, cattle, and performance. Also, it is thought that this study, which is the first bibliometric study in the field of animal science, will provide useful information to the researchers who will work in this field.

Key words: Bibliometric analysis, CiteSpace, animal science literature, social network analysis, citation analysis

1. Introduction

Bibliometric analysis is used to qualitatively and quantitatively analyze the effects of journals, institutions, research groups, individual researchers, or countries [1]. The use of bibliometric indicators has been increasing in recent years. Through bibliometric analysis and visualization, we can explore the intellectual landscape of a knowledge domain and discern what questions researchers have been trying to answer and which methods were used and developed for this purpose [2]. As a widely accepted definition in literature, bibliometric is the application of mathematical and statistical methods on articles, journals, and books [3,4]. In other words, bibliometric is defined as the numerical analysis of the publications produced by individuals or institutions in a given period and a specific topic area and the relations between these publications [5]. Bibliometric knowledge saves an ample amount of time for researchers to get started with the research of a domain and helps to inform about the major trends observed in the fields studied [6]. Bibliometric studies are of the nature of the studies carried out in the field and provide valuable information about the direction and quality of scientific researches [7]. Thus, it helps researchers to have an overview of the central studies and trending topics leading the field. The most widely used analysis in bibliometric studies are; author analysis, concept maps, cluster-factor analysis, citations, and reciprocal citations. In a sense, bibliometric studies do a citation network review in the background. For this reason, social network analysis, which is a useful tool for examining networks and the structures that make up the network, is accepted as a helpful method for bibliometric studies.

Social network analysis is an interdisciplinary research area built on the theoretical bases obtained from sociology, anthropology, statistics, mathematics, information sciences, education, psychology, and other disciplines over a long period [8]. Social network analysis aims to explain, visualize, and understand the network structure obtained from the relationships among individuals, objects, or units through statistical modeling. Many systems in nature and technology are examples of social networks [9]. Visual representation of social networks is quite significant in terms of understanding the data in the network and interpreting the results of the analysis more easily [10]. Most of the software developed for this purpose have various modules for the visualization of the network. The discovery of the data at hand, the display of the nodes and connections in different designs are realized by visualizing them in different shapes according to their colors, dimensions, and other advanced features.

Bibliometric analyses offer a useful tool to represent the available literature in a specific research field. It involves

* Correspondence: fatmayardibi@gmail.com

quantitative and visual processes to identify patterns and dynamics in scientific publications [11]. This study aims to determine trends in the animal science field by examining bibliometric features of the studies published in the literature, which are listed in WoS database. It was planned to discuss the knowledge structure and specific research themes in terms of leading researchers, authors, institutions, and countries in the field of animal science and their implications on the nature of animal science. In this way, the research will allow the researchers to have an overview of the intellectual structure and the current research themes in the field and reference books to be benefited from within the framework of the topic. The main question this study focused on was what were the authors, journals, countries of publication, subject fields that were effective in the works performed between 2015 and 2019 for animal science.

Thanks to CiteSpace software, in this study, the cocitation networks in the domain of animal science studies were described, the intellectual base and research front were noticed, and the critical points were analyzed. Thus, a general picture of animal science studies was presented. By defining intellectual milestones and dynamically visualizing citation networks, it will be easier to understand evolution, development, and trends in a particular scientific field.

2. Material and methods

In this study, the articles published between 2015 and 2019 on the animal science field were analyzed which were listed in Web of Science (WoS) database, and the last update of the database used for this study was 14.05.2019. Five journals that have the highest impact factor out of 60 journals in the Agriculture, Dairy, and Animal sciences category, according to the Incites Journal Citation Reports were selected for the study.

The number of publications that were studied was 6972 among these journals. The dataset was visualized and analyzed by using a new version of CiteSpace (5.4 R 4, issued on 2018).

2.1. Data analysis

As a result of the scanning performed on 6972 articles from the selected five journals in WoS Core Collection database, the following data were obtained: all bibliometric data, including the name of the authors and publications, title, source of the document, publication year, number of publications, number of citations, and type of the article were gathered together and saved as a text document. The academic works obtained after this operation were recorded into different data files, each containing 500 scientific works. The distribution of the works by years was shown in Table 1.

The literature type was defined as "all types" for the selected criteria. Six document types were found in these 6972 publications, and two publications were considered in the other category. The most frequent document type was the original article (n: 6605, 94.7%), accounting for 94.7% of total publications. The second position was review (226), with a proportion of 3.2%. The data was transferred to CiteSpace software for further analysis. CiteSpace is a Java application that combines information visualization methods, bibliometric, and data mining algorithms in an interactive visualization tool for the extraction of patterns in citation data [12]. CiteSpace software helps in finding the intellectual turning points and detecting burst terms, further demonstrating the dynamic changes and developments in animal science field studies and critical points in the development of a subject area or a discipline.

CiteSpace visualizes the networks in consequent years as a merged network. Dots in visualization represent the nodes in the networks. The lines connecting the nodes show the cocitation links and the color of the lines between nodes represents the year of citation that helps to understand which part of network is old and which part of network is new. The visualizations only illustrates highly cited publications, whereas CiteSpace lists the all citations in a table.

Table 1. Number of publications. The table shows that the total number of publications for each journal between 2015 and 2019.

	2019	2018	2017	2016	2015	Total
ARoAB	24	16	18	17	24	99
GSE	16	72	95	95	100	378
JoASaB	35	91	85	69	59	339
JoDS	430	1 024	932	955	853	4 194
PS	217	512	519	348	366	1 962

ARoAB: Annual Review of Animal Biosciences; GSE: Genetics Selection Evolution; JoASaB: Journal of Animal Science and Biotechnology; JoDS: Journal of Dairy Science; PS: Poultry Science.

3. Results and discussion

The number of citations that an article has was the main factor to reflect the quality of a paper. H-index, also known as H index or H factor (H-factor), stands for "high citations". According to the analysis of the data from WoS, the citations in all publications, and the H-index of different journals were respectively Annual Review of Anima Biosciences (ARoAB) 18, Genetics Selection Evolution (GSE) 22, Journal of Animal Science and Biotechnology (JoASaB) 20, Journal of Dairy Science (JoDS) 34 and Poultry Science (PS) 23. Although JoDS has the highest H-index value and sum of times cited value, it was found that ARoAB has the highest average citation per item. The higher the average citation per item value also affects the impact factor of the journal, in which AroAB has a higher impact factor than the other journals. The reason JoDS's sum of times cited value was higher than other journals was due to the fact that this journal has more publications each year than the other journals studied.

3.1. Analysis of data by reference

It was found that the 6972 articles that were analyzed in this study were cited 170,653 times. Figure 1 shows the network of the works with 30 or more citations. Articles that were cited less than 30 were not included in the network shown in Figure 1 since the threshold value was set as 30. The network shown in Figure 1 has a total of 187 nodes and 826 connections. The density of the network was found to be 0.0475. The thickness of a ring is comparable to the number of citations received in that time slice. Thus, a large-sized circle denotes a highly cited unit, reference.

The blue nodes in Figure 1 represent the studies that were published earlier (2015), while the yellow nodes represent article that was published recently (2019). The rings in Figure 1 depicted the citation history of a cited reference, with its thickness denoting its amount of citations within a time range. The larger the ring, the more article was cited. A line between two rings reflected the cocitation link of two cited references, with its thickness showing the strength of cocitation and its color showing the time of the first cooccurrence. The color bar on the top indicates different time slices of publication years (Blue: 2015, Purple: 2016, Pink: 2017, Orange: 2018, Yellow: 2019). Rings and lines with specific colors pertain to the corresponding time range. Also, the red color was usually used to label the citation burst, and the purple color was added to a ring to demonstrate a high betweenness centrality. [13] reported that betweenness centrality, measured according to the number of links passing a node in a network, implies the degree of significance of a node. Thus a node with high betweenness centrality value and citation frequency usually signifies a revolutionary scientific work that proposes new theories or innovations.

The modularity Q and the mean silhouette scores were two important metrics that tell us about the overall structural properties of the network. High modularity value means that there were secure connections among the nodes in the modules, but the relationship between the nodes of different modules is sparse [14,15]. The mean

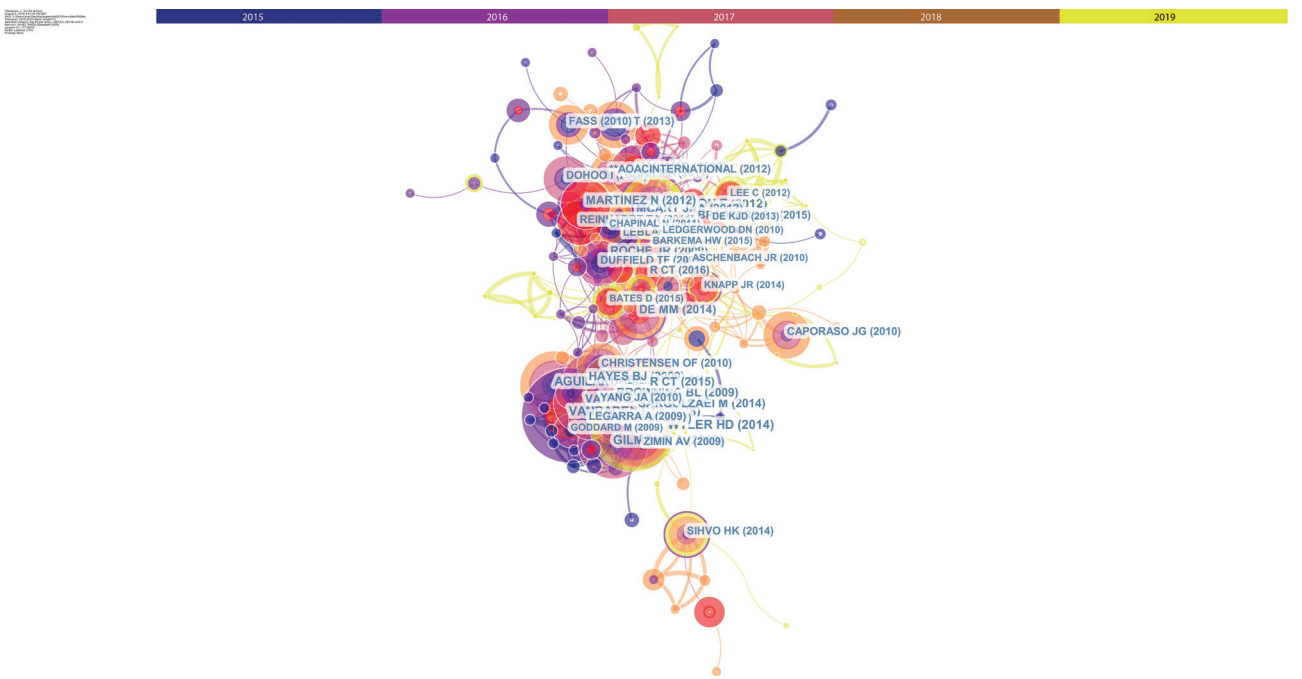

Figure 1. Visualization map of the reference network. The abbreviations that are written over the nodes correspond to the reference.

silhouette value, which indicates the similarity of the elements in a cluster, it is providing information about the structure of the clusters.

The modularity value was 0.6078, which means that the network is reasonably divided into loosely coupled clusters, and the mean silhouette value was 0.3277, suggesting that the homogeneity of these clusters was not very high. A total of 6972 academic works were divided into 17 clusters. Cluster analysis helps us understand the main features of science mapping [2]. CiteSpace provides different display modes, cluster view, and time-zone view. The time-zone view highlights the cocitation network changes with time, while the cluster view emphasizes the division of cocitation clusters within a period. In either time-zone view or cluster view of the cocitation network provided by CiteSpace, several critical attributes are represented by specified rings, lines, and colors.

Figure 2 shows a timeline visualization of the seven largest clusters and their interrelationships. Clusters were numbered from 0 to 6, cluster #0 (genomic prediction) was the largest cluster, and cluster #1 (bovine milk) was the second-largest one that was mentioned the most of the articles. It was observed from the results that the clusters have different durations. As shown in Figure 2, genomic prediction, bovine milk, and dairy calve clusters were sustained a long period of years, whereas the other clusters were relatively short-lived. Genomic prediction, bovine milk, and perinatal period clusters were active until 2018.

Since our study focused on these seven large clusters, the size and the silhouette values of these clusters were studied and shown in Table 2.

Cluster #0 was the largest cluster, containing 31 references across 11 years from 2006 till 2016. The median year of all references in this cluster was 2009. This cluster's silhouette value was 0.896. The silhouette column shows the homogeneity of a cluster. The higher the silhouette score, the more consistent of the cluster members were, provided the clusters in comparison have similar sizes [13].

Large-sized nodes or nodes with red tree rings were of particular interest in Figure 2 because they were either highly cited or have citation bursts or both. Thus, the result of burst analysis, which has been performed to see the most popular years of the works performed by different researchers, was shown that the highest citation burst value belongs that to the studies conducted by Vanraden PM, 2008.

3.2. Analysis of data by the author

The network has a total of 259 authors. JJ Loor was the author who had the highest productivity among others, with 73 times cited and belonged to cluster #3. Other highest cited authors were, SJ Leblanc (44 cited) belonged to cluster #1, and J Dijkstra (41 cited) belonged to cluster #0.

A total of 6972 academic studies were divided into 74 clusters, according to the authors. A timeline visualization

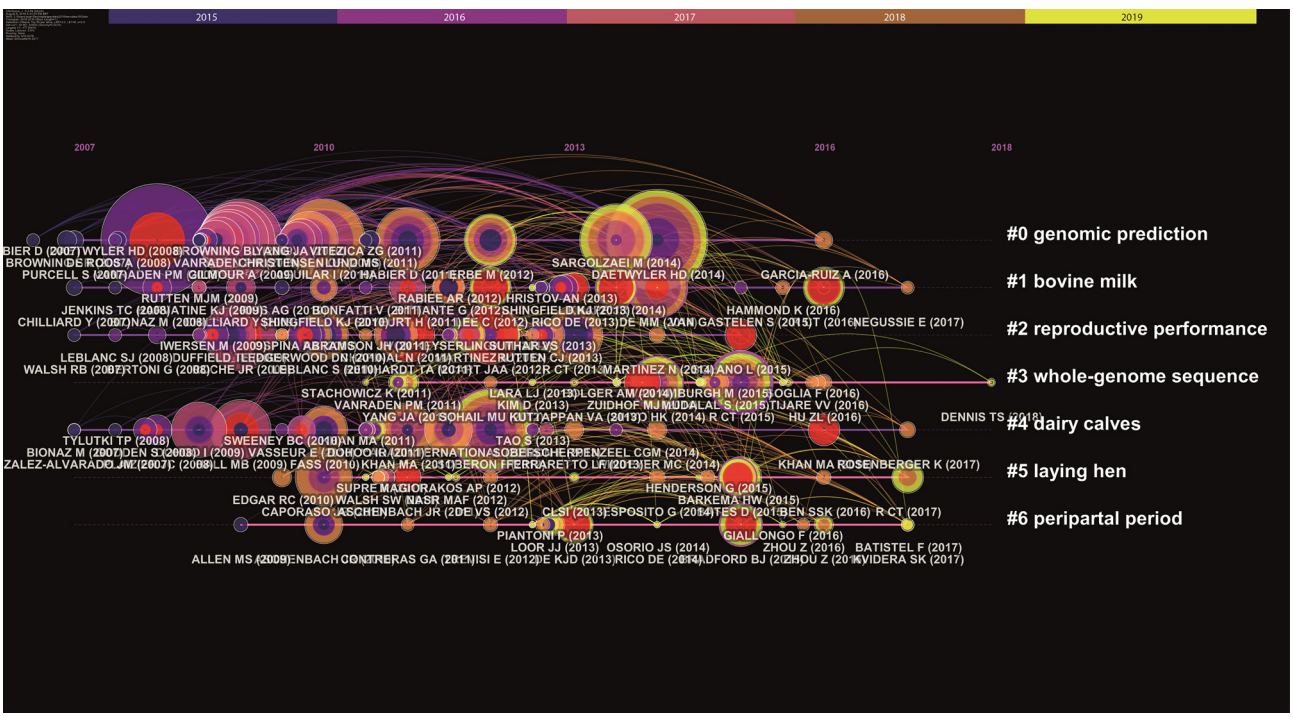

Figure 2. The references timeline view of the largest clusters. The abbreviations that are written over the nodes correspond to the reference.

Table 2. Summary of the clusters. The table shows that the cluster size, silhouette values, and mean (years) of the seven largest clusters automatically selected according to the reference.

Cluster ID	Size	Silhouette	Mean (year)
#0 genomic prediction	31	0.896	2009
#1 bovine milk	28	0.723	2012
#2 reproductive performance	27	0.754	2010
#3 whole-genome sequence	25	0.701	2014
4# dairy calve	24	0.850	2011
5# laying hen	22	0.775	2012
#6 peripartal period	20	0.872	2013

analysis of the largest clusters, according to the author, were cluster numbered from 0 to 11. Cluster #0 (rumen methane emission) was the largest cluster, and cluster #1 (commercial aviary) was the second-largest one.

As a result of the analysis, it was determined that a total of 259 authors were cited. Also, the results of burst analysis, which have been performed to see the most popular years of the works performed by different researchers, were shown that the author with the highest citation burst value is Nagendra P. Shan (2015), with 4.98. Citation burst, one of the most effective methods to determine research trends in the discipline, was occurred in Cluster #1 (commercial aviary), according to the author.

In timeline visualizations of cited authors, a cited author was positioned based on the earliest year in which he/she was cited in the dataset. A possible extension of this design would differentiate citations to the same author in different years [13].

A timeline visualization in Figure 3 shows that Clusters 3# (mammary epithelial cell), #6 (technological trait), and #11 (apparent ruminal synthesis) were not active clusters.

The homogeneity of the cluster is measured by silhouette value. If the cluster silhouette value is low, it is not shown in timeline visualization by the software due to cluster heterogeneity. Cluster labels, and the number of clusters are determined by the spectral clustering algorithm of CiteSpace software based on the optimal cut automatically, and the software does not allow analysts to determine the number of clusters there should be [13].

From the results, it has been seen that a specialization field has been developed over time, starting from the conceptualization stage. Moreover, some fields that have completed its development may have shifted to another area of expertise over time.

3.3. Analysis of data by country productivity

As a result of the country collaboration analysis, the network consisted of 68 nodes and 500 connections. The density of the country collaboration network was 0.2195. The country with the highest citation in animal science was the USA with 2595 citations; the country with the highest centrality was also the USA with a value of 0.18. Accordingly, it can be seen that the USA is already a well-known leader country. The country that has the second-highest citation was PRC with 1064 citations. On the other hand, the centrality values of Brazil and New Zealand were computed as 0.00 and it was concluded that both of them were not active in the animal science field. The citations of top ten countries are respectively listed as follows: USA (2595), China (1064), Canada (688), Brazil (427), Germany (393), Italy (389), Netherland (381), France (318), Australia (295) and Denmark (242).

3.4. Analysis of data by key words analysis

A key word analysis is an effective way to show emerging trends and track research topics over time because key words provide a concise summary of a document. Key word analysis was performed without any restrictions based on the frequency of the words. The network that has been formed accordingly has 113 nodes and 719 connections. The density of the network was found to be 0.1136. The number of repetitions of each word, the years of repetition and the centrality values of the words were displayed in Table 3, allowing statistical interpretation of visual results.

The modularity value was 0.3997, and the mean silhouette value was 0.6191. According to mean silhouette value, which indicates the similarity of the elements in a cluster, it was observed that academic works included within the extent of the study are well-clustered. The silhouette value close to 1 indicates functional clustering. The network was divided into a total of 5 clusters, as displayed in Figure 4. As shown in Figure 4, the largest cluster was growth performance.

When the top 10 key words were reviewed, "performance", "broiler", "chicken", "growth", "growth performance" belonged to the highest cluster #0 (growth performance), and it was still an active cluster.

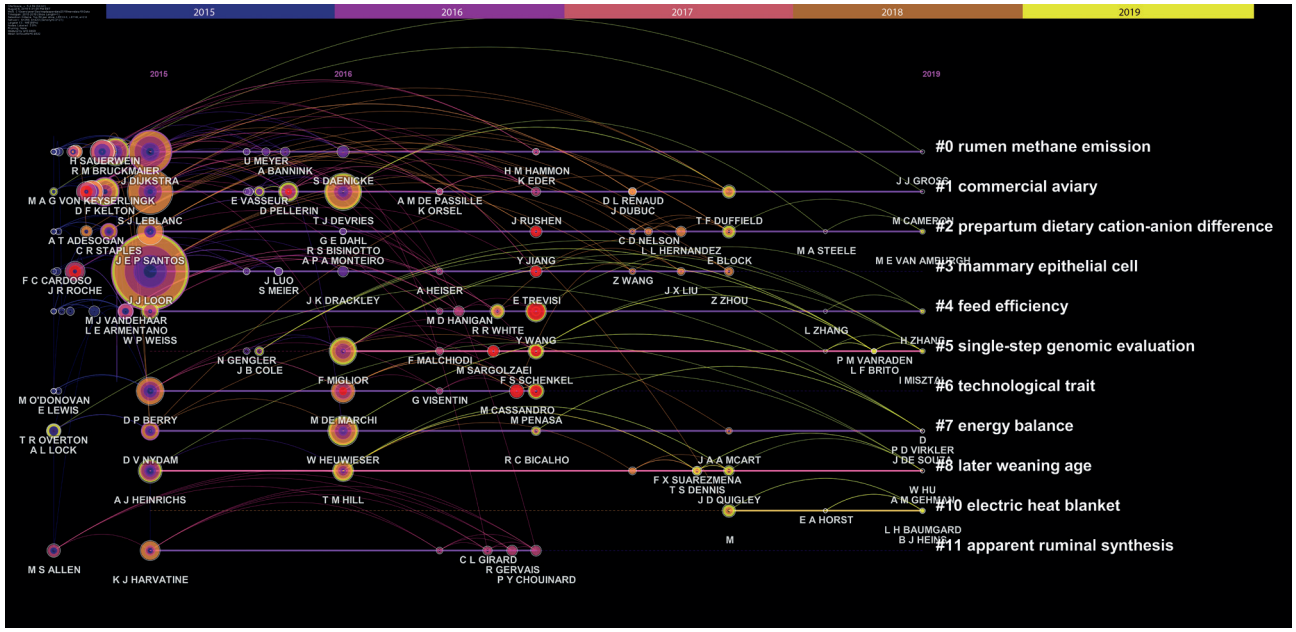

Figure 3. The authors timeline view of the largest clusters. The abbreviations that are written over the nodes correspond to author names.

Table 3. Top 10 key words. The table shows that the number of repetitions of each key word and the clusters to which they belong, the years of repetition of the key word, and the centrality values.

Frequency	Centrality	Year	Cluster	Key words
942	0.16	2015	#2	dairy cow
747	0.09	2015	#4	cattle
686	0.08	2015	#0	performance
493	0.08	2015	#0	broiler
447	0.08	2015	#0	chicken
447	0.11	2015	#2	milk production
414	0.08	2015	#0	growth
398	0.08	2015	#4	cow
389	0.02	2015	#0	growth performance
385	0.07	2015	#1	milk

As can be seen from Figure 4, the largest cluster (#0 growth performance) and the second-largest cluster #1 (staphylococcus) were still active clusters. In contrast, the #4 (technical note) cluster had been ended by 2017.

Five highest citation burst value was displayed in Table 4. As shown in the table, "nutrient digestibility" was defined as an active keyword since 2017, and this may become a popular topic in the future. Besides, citation bursts were mostly cluster #2. Thus we can say that staphylococcus aureus is an active field.

3.5. Analysis of data by productivity of institution

More than 50 institutions were identified in 6971 works. Table 1 was created according to the results of the productivity analysis of the institutions. The network that has been formed accordingly has 87 nodes and 595 connections. The density of the network was found to be 0.159.

The analysis showed that the critical publications came from University of Guelph. At the same time, University of Guelph is a high central institution in this area. The modularity value was 0.2763, and the mean silhouette value was 0.4159. The network was divided into a total of 7 clusters. The largest and the second-largest clusters were quantitative trait loci and genomic prediction, respectively.

As a result of the timeline visualization analysis, the largest cluster (#0 quantitative trait loci) was not an active

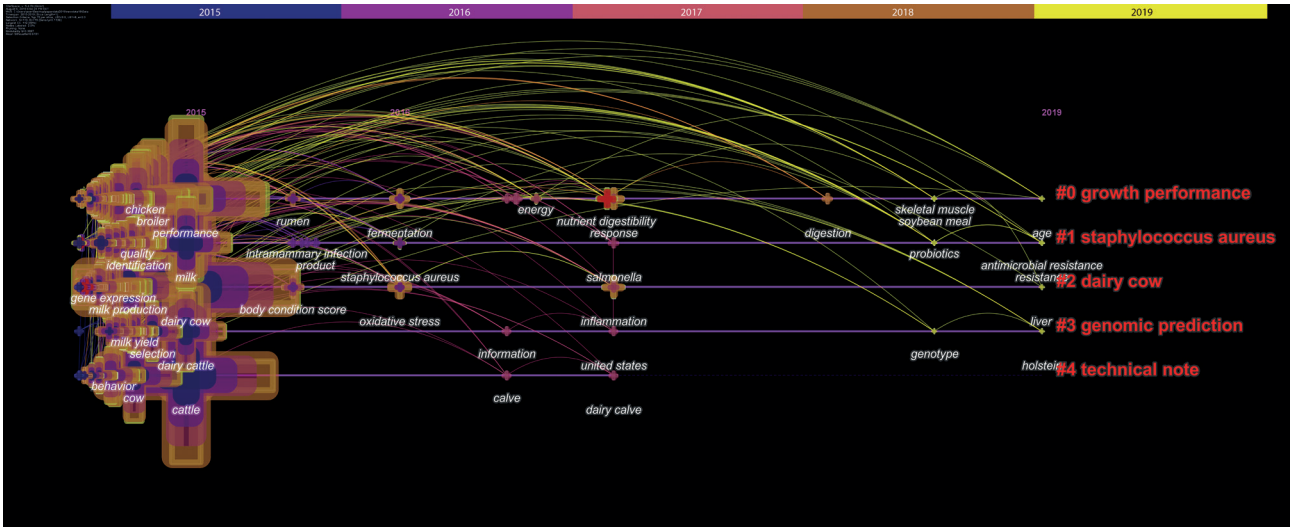

Figure 4. The key words timeline view of the largest clusters.

Table 4. Top 5 key words with the strongest citation bursts. The table shows that the top 5 key words with the highest citation bursts and their years of popularity. The red color bar in the last column of the table represents the time period of the cited years. Total length of two different color bar represents the timeline starting from the year 2015 to the year 2019.

Key words	Year	Cluster	Strength	Begin	End	2015–2019
bacteria	2015	#1	18.5069	2015	2016	
energy balance	2015	#2	16.2194	2015	2016	
sheep	2015	#2	4.7603	2015	2016	
conjugated linoleic acid	2015	#2	3.8222	2015	2017	
nutrient digestibility	2015	#0	9.0056	2017	2019	

cluster. This cluster was active during the period between 2015 and 2017. Second-largest cluster #1 (genomic prediction) and #2 (lactobacillus plantarum zdy2013) were still active clusters. Moreover, cluster #2 mean silhouette score of cluster # 2 was 0.534, which suggests that the homogeneity of this cluster on average was not very high, but not very low either. Top cited references in this cluster were mostly articles published in early 2015.

Wageningen University was the most active institution with the highest burst, and it belonged to cluster #1 (genomic prediction).

4. Conclusion

This study has been performed to determine the outstanding authors, journals, countries, and subject areas of the works completed in the animal science area through a bibliometric analysis of the academic works published in the WoS database. A total of 6972 academic works published between 2015 and 2019 have been analyzed through CiteSpace software, and the outcomes were reported both graphically and statistically. The evidence revealed in this study suggests that 'genomic prediction' is the most effective topic of study in the animal science field and will be still an active topic in the future. The most effective country is the USA, which is the most central state of the domain, and many critical publications originated from the USA.

To avoid misunderstanding and to guide future investigations, the main limitation of this study should be noted. Even though the data covered the most critical articles in the field of animal science in the WoS database by CiteSpace comprehensively, some of the essential journals were excluded. It is thought that this study will set an example for scientists studying animal science for their future studies in terms of the performance of the analysis.

5. Software and data repository resources

CiteSpace software can be downloaded from http://cluster.cis.drexel.edu/~cchen/citespace/. All data used in this study are from the Web of Science database.

References

1. Kamdem JP, Duarte AE, Lima KRR, Rocha JBT, Hassan W et al. Research Trends in Food Chemistry: A bibliometric Review of its 40 years Anniversary (1976-2016). Food chemistry 2019; 294: 448-457. doi: 10.1016/j.foodchem.2019.05.021
2. Chen C. CiteSpace: A Practical Guide for Mapping Scientific Literature. NY, USA: Nova Science Publishers; 2016.
3. Pritchard A. Statistical bibliography or bibliometrics. Journal of Documentation 1969; 25 (4): 348-349.
4. Chen D, Liu Z, Luo Z, Webber M, Chen J. Bibliometric and visualized analysis of emergy research. Ecological Engineering 2016; 90: 285-293. doi: 10.1016/j.ecoleng.2016.01.026
5. Ellegaard O, Wallin JA. The bibliometric analysis of scholarly production: How great is the impact? Scientometrics 2015; 105 (2015): 1809-1831. doi: 10.1007/s11192-015-1645-z
6. Iqbal W, Qadir J, Tyson G, Mian AN, Hassan S-u, et al. A bibliometric analysis of publications in computer networking research. Scientometrics 2019; 119 (2): 1121-1155. doi: 10.1007/s11192-019-03086-z
7. Abejón R, Pérez-Acebo H, Garea A. A bibliometric analysis of research on supported ionic liquid membranes during the 1995–2015 period: Study of the main applications and trending topics. Membranes 2017; 7 (4): 63. doi: 10.3390/membranes7040063.
8. Van Duijn MA, Vermunt JK. What is special about social network analysis? Methodology 2006; 2 (1): 2-6. doi: 10.1027/1614-2241.2.1.2
9. Crossley N, Prell C, Scott J. Social network analysis: introduction to special edition. Methodological Innovations Online 2009; 4 (1): 1-5. doi: 10.1177/205979910900400101
10. Hogan B, Carrasco JA, Wellman B. Visualizing personal networks: Working with participant-aided sociograms. Field Methods 2007; 19 (2): 116-144. doi: 10.1177/1525822X06298589
11. Yagci MI, Kuhzady S, Balik ZS, Öztürk L. In Search of Consumer's Black Box: A Bibliometric Analysis of Neuromarketing Research. Journal of Consumer and Consumption Research 2018; 10 (1): 101-134.
12. Synnestvedt MB, Chen C, Holmes JH. CiteSpace II: visualization and knowledge discovery in bibliographic databases. AMIA Annual Symposium Proceedings 2005; 724-728.
13. Chen C, Ibekwe-SanJuan F, Hou J. The structure and dynamics of cocitation clusters: A multiple-perspective cocitation analysis. Journal of the American Society for Information Science and Technology 2010; 61 (7): 1386-1409. doi: 10.1002/asi.21309
14. Li X, Ma E, Qu H. Knowledge mapping of hospitality research–A visual analysis using CiteSpace. International Journal of Hospitality Management 2017; 60: 77-93. doi: 10.1016/j.ijhm.2016.10.006.
15. Guo S, Wang L, Xie Y, Luo X, Zhang S, et al. Bibliometric and visualized analysis of stem cells therapy for spinal cord injury based on Web of Science and CiteSpace in the last 20 years. World neurosurgery 2019; 132: e246-e258. doi: 10.1016/j.wneu.2019.08.191