2014 Proceedings of PICMET '14: Infrastructure and Service Integration.

Study on the Academic Landscape of Hydropower: A Citation-Analysis Based Method and Its Application

Zhidong Liu, Hajime Sasaki, Ichiro Sakata
Dept. of Technology Management for Innovation, Graduate School of Engineering, The University of Tokyo, Tokyo, Japan

Abstract—**Hydropower, with characteristics of superior storage capacity and speedy response, can meet sudden fluctuations electricity demand. And its generation capacity could be doubled by 2050 according to IEA's recent estimation. Besides, hydropower doesn't require innovative technologies, and can be operated with minimal cost. However, no big picture of knowledge related to hydropower can be found in the preceding research. This paper aims to figure out the academic landscape in this field by analyzing the citation network of papers published in academic journals. We collected 7,521 target papers using specific search query from Web of Science. By utilizing a topological based method, all the papers were categorized into clusters by their own characteristic topics. Results show the existence of 6 principal research clusters: Renewable energy, Optimization of system operation, Environmental issues (greenhouse-gas emission), Environmental issues (fish management), Environmental issues (sediment) and Pumped hydro storage systems. As the keywords of some major clusters seem to be general, in-depth sub-clusters analysis was also conducted to gain better knowledge of those clusters. Combing the analysis results, (sub)clusters related to pumped-storage and small hydropower are considered to be developing, as indicated by the average publication year of papers and recent increasing trend.**

I. INTRODUCTION

Hydropower or more precisely speaking hydroelectricity as discussed in this study is the term referring to electricity derived from the potential energy of water captured when moving from higher to lower elevations. Categories of hydropower projects include run-of-river, reservoir-based capacity, and low-head in-stream technology (the least developed). Hydropower covers a continuum in project scale from large (usually defined as more than 10 MW installed capacity, while the definition varies by country) to small, mini, micro, and pico (widely used in rural electrification). Regardless of the rapid development of other renewable energy, it still remains by far the largest source of renewable energy worldwide, accounting for one fifth of the world's electricity supply [1]. Being technology whose concept was first put forward by a French engineer in the mid-1770s, hydropower as an interdisciplinary research field is still enjoying popularity even now. As shown in Fig. 1, the number of yearly publications is increasing exponentially, and the accumulated number of publications is also increasing exponentially.

With thousands of academic papers being published on this subject annually, the fast-changing and complicated situation has made it impossible for researchers even those experts in this domain to understand the whole picture by reading through all the papers. Empirical knowledge, case studies and brief surveys are no longer sufficient, where grasping the current status of hydropower has become an imperative task.

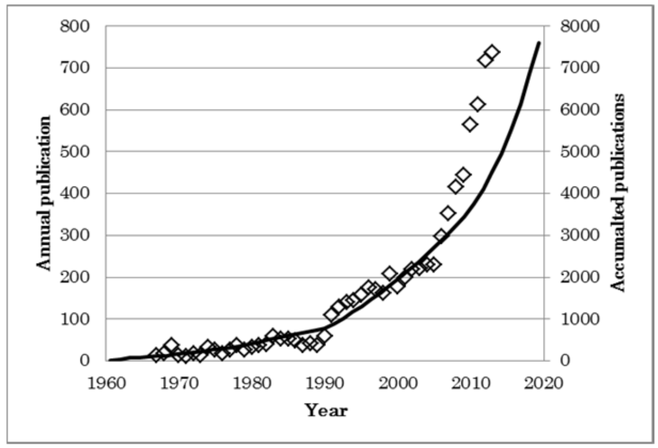

Fig. 1 Number of papers retrieved from database. White square and line represent the number of annual publications and the accumulated numbers of publications, respectively

To meet the challenge caused by the huge scale of information, a computer-based approach can be introduced to complement the expert-based approach [2] [3]. This approach bases on an understandable assumption that citing and cited papers have similar or at least related research topics. And the interrelationships among thousands of papers make up a complex citation network. By analyzing the citation work, we can comprehend the structure of a research domain consisting of a large volume of academic papers without reading through all of them.

After obtaining the academic landscape of hydropower, this research can benefit the society in several ways. First of all, for the academia side, researchers in this domain can benefit from the fundamental framework of the current status in this research, who will also no longer need to overread such large number of papers. It would become quite easy to locate the trendy research topic indicated by groups of newly published papers around the center topics. While, for the industry side, hydropower itself as a renewable energy with lower technology requirement and lower cost can be found operating almost all around the world. Recent years, with the rapid economic growth of emerging countries, how to meet the increasing energy demand with limited budget has become a serious social problem or even political problems in some cases. Our research can also contribute to the technology acquisition roadmap for those emerging countries. With the academic landscape, energy planners can have an instant image of current status and determine the general direction. They can also find better alternative methods recently available or those can be expected in the near future. In this way, the distance between academia and industry can be shortened. Ultimately, the society as a whole can benefit from our research.

As for the existing papers, since our research is related to the literature review of hydropower and citation network analysis, we studied about the previous works from both sides. On one side, as for the literature review under the topic of hydropower, most of them usually concentrate on certain topics under hydropower, like small hydropower [1] [4], hydropower plant operation [5] [6], pumped-hydro energy storage [7]. Most of these papers discuss about the current progress in the research of hydropower related by listing up and classifying various technologies with detailed analysis. However, they may not be able to provide a big picture of research on hydropower since they tend to focus on some trendy topics while omitting others. Also, some papers talk about present situation and future prospect of hydropower in certain countries or areas, which aslo include discussion covering a series of proper technologies accoding to local conditions [8]. But they cannot provide analysis on a global scale, and nothing can be learned about the development of technologies happening in other countries at the same time. On the other side, being a comprehensive analysis tool, a citation-based approach has been applied tostudies of various fields, like that in nanotechnology [9] and water resource management [10], of course some preceding studies have been done related to other renewable energy like solar power [11]. However, up to now, according to the literature scan we have conducted, there seems to be no similar researches having been done about hydropower. This also contributes to our decision to undertake this research.

Back to our research, the major objective of this paper is to figure out the academic landscape of researches in the field of hydropower, which is seldom done in the prior research as mentioned above. And based on the academic landscape of contemporary status, a forecast of the future development in the field of hydropower can be made.

II. METHOD

In order to conduct the citation network analysis, we collected a set of academic papers related to hydropower, which are defined in this as including “hydropower” or “hydroelectricity” in their titles, abstracts and keywords. We obtained citation data of hydropower-related publications from the Science Citation Index (SCI) and the Social Science Citation Index (SSCI) compiled by the Institute for Science Information (ISI), since both of them are the best sources for citation data. We used the Web of Science to access to these databases, which is an online subscription-based scientific citation indexing service maintained by Thomson Reuters. First of all, we searched the papers using the query of *(("power" or electric*) near "hydro") or hydro-power or hydropower or hydroelectric**. The asterisk * represents a wildcard, and by using specific search operators “NEAR”, the system can be told to find records where the terms joined by NEAR are within 15 words of each other. Both of them can help to locate the right results. A total of 7521 papers were retrieved, which contain the keyword that we are interested in. However, not all of the retrieved papers are closely relevant to the major topic of hydropower, or they can be called noise. Therefore, we focused on the maximum connected component, which account for about 43.2% of the total papers (3250 out of 7521 papers). It means that we regarded papers having no citation relation with other papers as digressional and will not be taken into consideration in this research. Then, the citation network is converted to a non-weighted, non-directed network as shown in Fig. 2. Finally, by using topological clustering method, the network is divided into several clusters [12] [13].

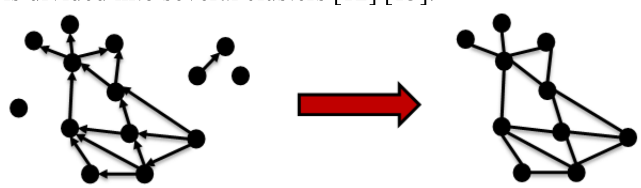

Fig. 2 Schematics of the process to convert the retrieved data into maximum connected component

After clustering the network, we analyzed the characteristics of each cluster by the titles and abstracts of papers that are frequently cited by the other papers in the cluster, as well as the journals in which the papers in the cluster were published. We named each cluster and also listed the keywords for each cluster from the titles and abstracts of the top twenty most cited papers in the cluster. The average publication year of papers in each cluster was calculated to know the trendy research field in hydropower research, which may lead to emerging technologies in the near future.

III. RESULTS

By entering the search query into the search engine of Web of Science, we successfully retrieved 7521 related papers from the whole database (up to 2013/9/29). As shown in Fig. 3, as an academic field with a long history, the earliest publication recorded in the database can date back to 1908. Different from the slow increase in the early years, there exists a sharp increase of annual publication starting from around 1990s, which is just the same time sustainable development as a concept was proposed in *Our Common Future* by the World Commission on Environment and Development (WCED) [14]. Because of the lack of sufficient publication data of 2013, there is a decline in the number of academic papers between 2012 and 2013.

After the process of clusterization, the whole citation network of hydropower can be divided into 50 clusters, where the number of nodes in each cluster varies with each other with the largest cluster of 656 nodes and the smallest one containing only 3 nodes. As explained in the last chapter, papers in each cluster are strongly coupled by intra-cluster citation. As indicated in Fig. 4, the number of nodes in each cluster decreases sharply until the 9th cluster, and after the 20th cluster the number becomes trivial. In this way, we can neglect the clusters after the No.6 clusters, since the top 6 clusters account for about 80% of the total papers in this network as shown in Fig. 4.

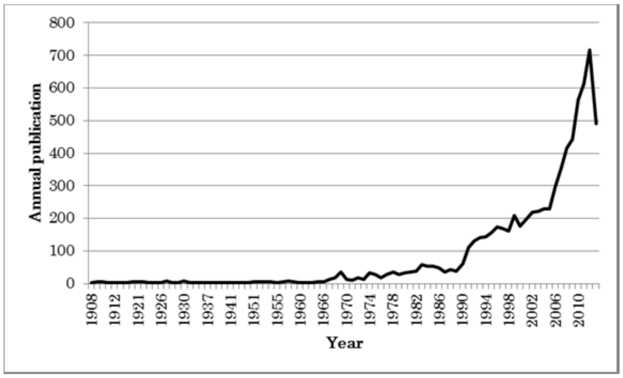

Fig. 3 Time series analysis of the retrieved papers

By assigning different colors to intra-cluster links in different clusters, we can obtain the visualized structure of the citation work, which is Fig. 5. You may notice some clusters' structures are compact and round like Cluster #2, while some others are stretched like Cluster #3. Usually, the former ones stand for clusters that consist of papers with strong tendency to cite other papers in the same cluster, while the latter ones are closely related to the clusters located in corresponding direction. When clusters are close to each other, it means that there exist close citation relations among papers in those clusters. Table 1 summarizes the characteristics of the top 6 clusters.

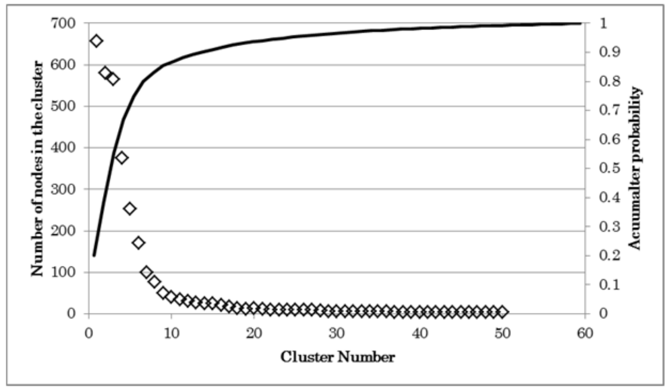

Fig. 4 Cluster size. White square and line represent the number of nodes in each cluster and the cumulative probability of the number of nodes, respectively.

As shown in Table 1, Cluster #1 has 656 papers in it and is the biggest among the 50 clusters. And, the top keywords of Cluster #1 turn out to be quite general and a little bit different, as the top 1 keyword is a nation's name "turkey", which seems to be not so closely related to the central topic of hydropower. In order to know the contents of this cluster, we further conduct an in-depth clusterization of the papers in Cluster #1. And the visualized result is shown as Fig. 6. The top 3 sub-clusters are about hydropower in Turkey, small hydropower, and the effect of climate change on hydropower. Despite the fact that "turkey" ranks No.1 among all the keywords of Cluster #1, the biggest sub-cluster is sub-cluster #1.1, which is about small hydropower. It is also the youngest one among the sub-clusters of Cluster #1, corresponding to the current trend of exploiting small hydropower. Still, we cannot ignore the fact that Turkish researchers have contributed the most to the study in this field according to the country rank. According to a recent report, Turkey is

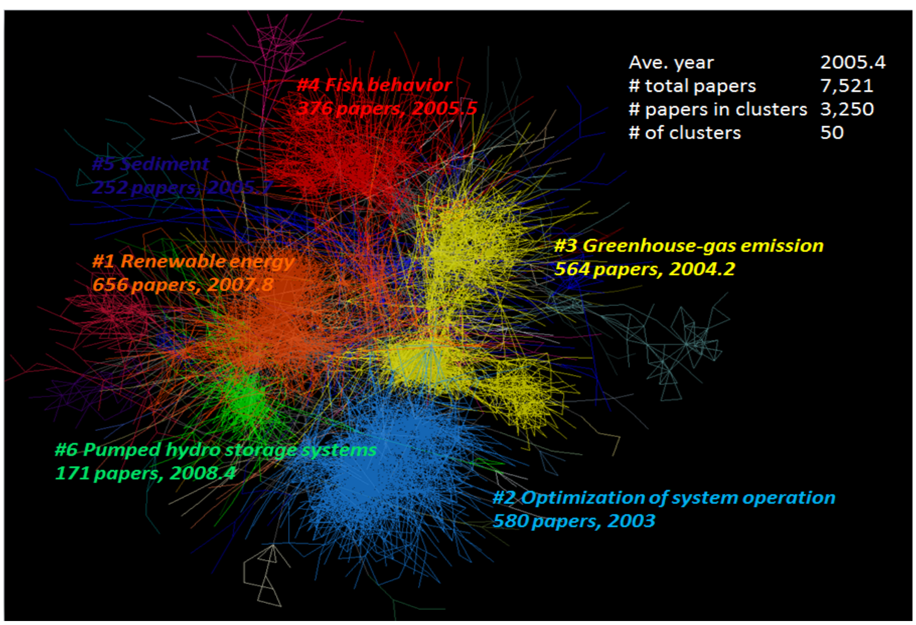

Fig. 5 Visualization of the structure of hydropower

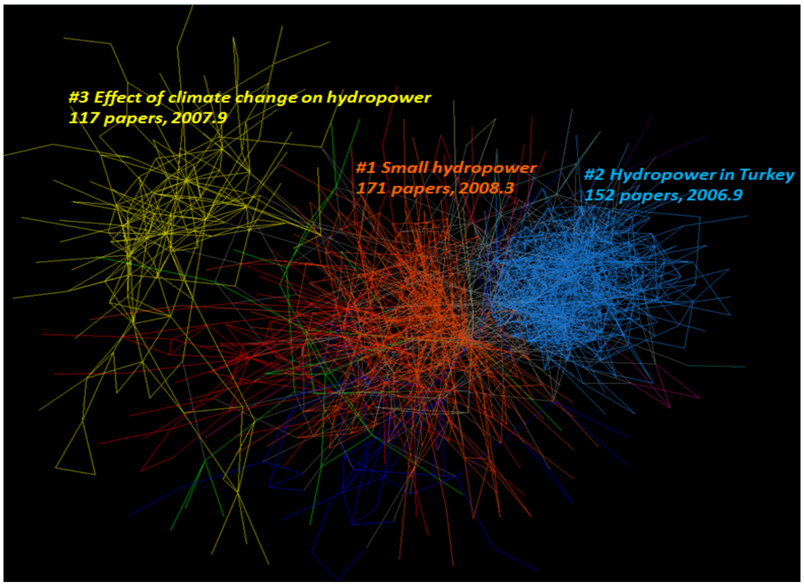

Fig. 6 Visualization of the sub-clusters of Cluster #1

increasing its hydropower capacity at a rapid rate to address chronic shortages of electricity and frequent power outages, which is approximately 2 GW in 2012, only second to China [15]. Currently, hydropower is Turkey's major renewable energy source supplying approximately 25.0% of the country's annual electricity generation and the Turkish government has a very impressive target to utilise all the technically available hydroelectric potential by 2023 [16]. We think the market needs for hydroelectricity contribute a lot to the prosperity of related research in Turkey.

Cluster #2 is mainly about the method and algorithm to maximize the efficiency of the hydroelectricity generation, which is dominated by the United States. Judging from the rank of main journal, it is more inclined to technical issues rather than social issues in other major clusters. Besides, Cluster #2 is also the oldest cluster among the major clusters. As for the next 3 clusters, they mainly deal with environmental issues directly associated with the construction of hydroelectric reservoirs, like greenhouse-gas emission, influence on fish behavior, sediment problem. Since Cluster #2 to #4 are all closely related to practical development of hydropower, these 4 clusters are dominated by China, Brazil, Canada and the United States, which are also the top 4 hydropower generating countries in the world [15].

As for Cluster #6, it is the youngest major cluster, in which wind, pumped storage are the key topics. Wind-hydro, Solar-hydro, and other similar types of hybrid renewable energy system are discussed in this cluster. Unlike hydropower, generation of wind or solar energy bears stochastic nature, simultaneously the process of consumption is also uneven and depends on the rhythm of people’s life, type of energy consumption etc. For the agreement of demand and supply, there is a possibility of the storage of the part of the energy in the times of surplus and the energy output in the times of deficit, where hydropower just meet this kind of requirement. Rather than the United States, researchers from Croatia and Greece contribute the most papers to this cluster.

TABLE 1 CHARACTERISTICS OF THE TOP 6 CLUSTERS

No.	Cluster name	#Node	Average year	Main country	Main Institute	Main journal	Top keywords
#1	Renewable energy	656	2007.8	Turkey	Karadeniz Tech Univ (Turkey)	RENEW SUST ENERG REV	turkey
				USA	Sakarya Univ (Turkey)	RENEW ENERG	renewable
				India	Univ Washington (USA)	ENERG POLICY	renewable energy
#2	Optimization of system operation	580	2003	USA	Huazhong Univ Sci & Technol (China)	IEEE T POWER SYST	programming
				China	Iran Univ Sci & Technol (Iran)	J WATER RES PL-ASCE	scheduling
				Brazil	Univ Calif Davis (USA)	WATER RESOUR MANAG	optimization
#3	Greenhouse-gas emission	564	2004.2	Canada	Univ Quebec (Canada)	RIVER RES APPL	fish
				USA	Fisheries & Oceans Canada (Canada)	SCI TOTAL ENVIRON	mercury
				Brazil	Chinese Acad Sci (China)	REGUL RIVER	mehg
#4	Fish management	376	2005.5	USA	Pacific Nw Natl Lab (USA)	T AM FISH SOC	fish
				Canada	Univ Washington (USA)	N AM J FISH MANAGE	salmon
				Brazil	Us Geol Survey (USA)	HYDROBIOLOGIA	migration
#5	Sediment	252	2005.7	USA	Yunnan Univ (China)	HYDROBIOLOGIA	sediment
				Canada	Beijing Normal Univ (China)	GEOMORPHOLOGY	ice
				China	Univ British Columbia (Canada)	REGUL RIVER	lake
#6	Pumped hydro storage systems	171	2008.4	Croatia	Univ Zagreb (Croatia)	ENERG POLICY	wind
				USA	Univ Split (Croatia)	RENEW ENERG	storage
				Greece	Natl Tech Univ Athens (Greece)	APPL ENERG	pumped

IV. DISCUSSION

By plotting the annual publication data (1970-2013) of each cluster in one graph, we can obtain Fig. 7 below. Although varying from each other, the top 6 major clusters still share several traits in common: First of all, there are few papers being published before 1980s, indicating current dominating research topics are still young compared with the long history of hydropower. Also there is a rising tendency in recent years, this trend might be explained by the rapid development of research on sustainability science, which has become a significant issue in contemporary society [17]. Hydropower is no longer simply a method to produce electricity, more and more attention has been paid to other related studies, such as environmental impact of hydropower in Cluster #3, #4 and #5 [18]. And it seems like that the growth trend can be classified into three types. One is the rapid increase after 2008 as seen in Cluster #1 and #6. The other is the smooth increase as shown in Cluster #3, #4 and #5. The third is a fluctuation in the quantity of academic publication as in Cluster #2, for example, the number of publications in Cluster #2 drops from the late 1990s and increases again in the late 2000s.

As mentioned earlier, sub-cluster #1.1 and Cluster #6 are definitely young and rapidly developing among the major clusters. As far as we concerned, the appearance and development of these two topics is closely related to contemporary social needs. In the case of sub-cluster #1.1, hydropower on a small-scale is not only one of the most cost-effective energy technologies appropriate for rural electrification in developing countries, but also the main prospect for future hydro developments in Europe, where the large-scale opportunities have either been exploited already, or would now be considered environmentally unacceptable [1]. How to make electricity generation more efficiently and how to strike a balance to maintain both the health of the stream and the economics are the main topics of this research field. When it comes to Cluster #6, where hydropower acts as a large energy container to stabilize the output of electricity. Although the concept of pumped-storage hydroelectricity can date back to 1930s, it is the rapid development of wind, solar and other renewable energy that lead to the research and application of pumped-storage system in wind or solar power plant. The following Fig. 8 is based on wind electricity generation data from U.S. Energy Information Administration [19]. We can see from the chart that for most countries listed here, the growth starts around the year of 2000 and then continues at an almost constant increasing rate. Especially for United States and China, the growth accelerated in 2007 and has been fast increasing ever since. Compared with hydropower, wind and solar power are relatively younger. We consider part of the reason for Cluster #6 being the youngest cluster is to do with the progress of other renewable energy.

An estimated 30 GW of new hydropower capacity came on line in 2012, increasing global installed capacity by about 3% to an estimated 990 GW. The top countries for hydro capacity are China, Brazil, the United States, Canada, and Russia, which together account for 52% of total installed capacity [15]. We managed to obtain the hydroelectricity generation data of 2011 [19]. By combining it with the academic publication data, we can obtain Table 2 and Fig. 9 indicating the tight relation between industy and academia sides.

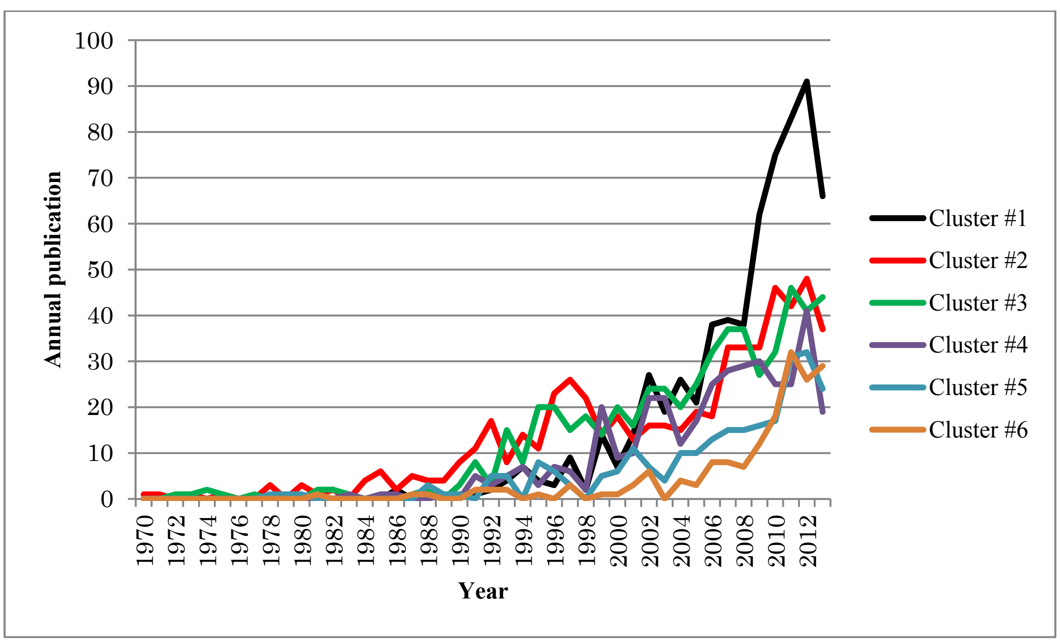

Fig. 7 Time series analyses of each cluster

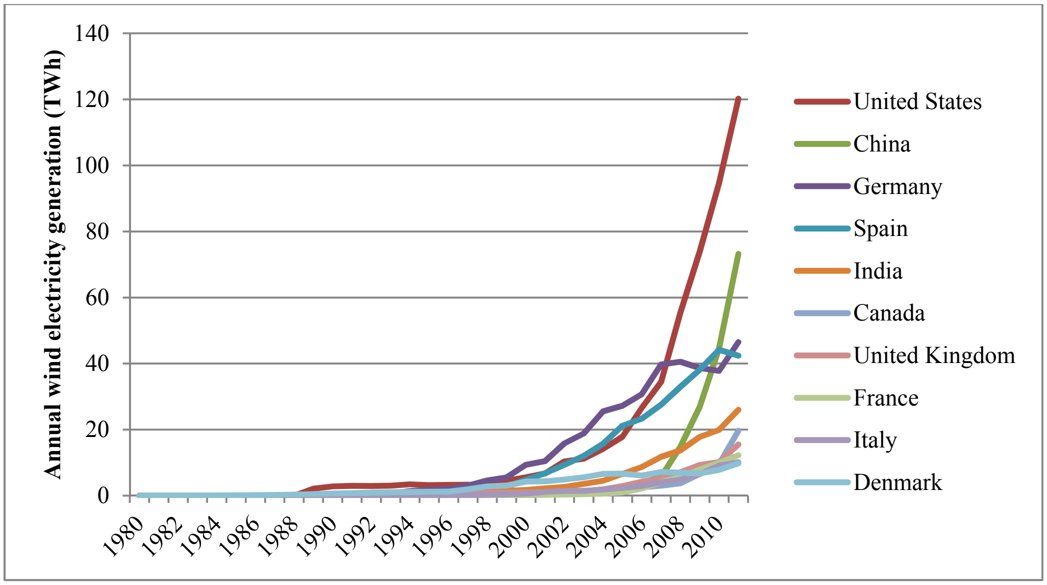

Fig. 8 Top 10 wind electricity countries (1980-2011)

TABLE 2 HYDROELECTRICITY GENERATION AND ACADEMIC PUBLICATION OF LEADING COUNTRIES

Country	Hydroelectricity generation (2011)		Academic Publication	
	Rank	Generation	Rank	Publications
China	1	694	3	252
Brazil	2	430	4	239
Canada	3	377	2	316
United States	4	328	1	706
Russia	5	165	**50**	7
India	6	132	6	170
Norway	7	122	12	84
Japan	8	85	24	33
Venezuela	9	84	**58**	6
Sweden	10	66	14	75

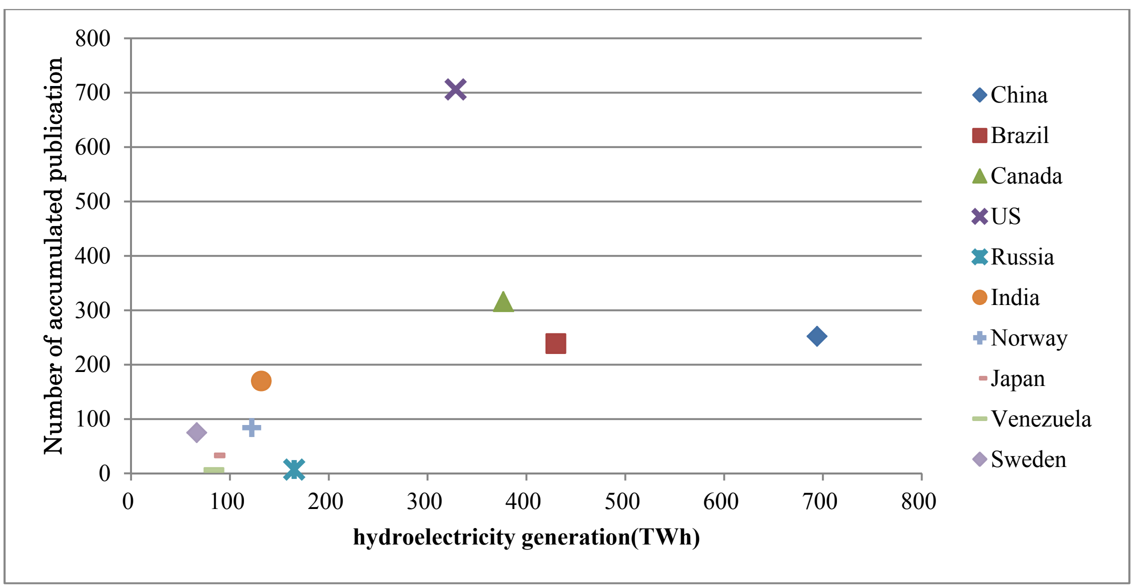

Fig. 9 scatter chart showing hydroelectricity generation and academic publication of leading countries

As shown in Table 2 and Fig. 9, except for Russia and Venezuela, countries with large electricity generation also possess a high rank in the academic publication. Furthermore, the top 4 countries of both ranks are the same, despite some trivial difference in the order, which is subject to many other factors like natural environment. Therefore, we consider that the relation between industry and academia is very close in the case of hydropower.

In the end, let us address the limitations of the approach applied in this research. We collected the corpus of academic publications by making a query, where we may fail to include some related papers or exclude unrelated papers. The results obtained by citation network analysis indicate that the top 6 major clusters account for the majority of the total papers, among which small hydropower and pumped-storage system are the emerging research fields with an acute increase in recent years. Besides, many papers are discussing about the social and environmental influence that this technology has brought or will possibly bring to the society. It can be explained that as a mature technology the room left for new breakthrough is limited. Instead of putting all the research focus on technology issues, more and more researchers are now dealing with non-technology issues related to hydropower. However, even we can prove that there exists a close relation between academic research and practical application, it does not mean that all the scientific progress can be commercialized to meet the social needs. The bibliographic analysis may not be able to offer enough grounds for making a certain prediction of how emerging technology will evolve without information on the application side. Still, the citation network approach is a powerful tool to assist experts to construct fundamental frameworks in domains where the quantity of annual publication is far beyond handling ability just like hydropower.

V. CONCLUDING REMARKS

Hydropower as one of the most important renewable energy is indispensable for the future development of our society and economics. In order to meet the rapidly increasing needs of electricity, a growing body of research efforts on hydropower can be observed. Despite the fact that no one can read through all the papers being published every year, planners of energy research or energy researchers still need to grasp the broader coverage of hydropower research, and make decisions of effective investment with limited resources. The computer-based approach is expected to offer supplemental information.

This paper visualized the structure of hydropower by analysis based on citations relation among relevant publications, and used a topological clustering method to detect the major topics of hydropower. Our citation analysis extracted 6 main research domains: Renewable energy, Optimization of system operation, Environmental issues (greenhouse-gas emission), Environmental issues (fish management), Environmental issues (sediment) and Pumped hydro storage systems. Among these major clusters, small hydropower and pumped-storage system are currently developing, which is corresponding to the development of small hydropower and hybrid power plants on the application side. By comparing the data of hydroelectricity generation and academic publication, we consider that the relation between academia and industry of hydropower is quite close and tight.

The analysis can offer a concise summary of the overall structure of the target research field and emerging research topics there. However, we discarded the most of the contents of the papers during analysis. For instance, each cluster is named after the title and abstracts of the top twenty most cited papers in the cluster. In this way, information from less-cited papers is ignored, which in some situations might have great potential as novel technology. Besides, our analysis is greatly determined by the pile of publications we extracted in the first step. It is not an easy task to define a research domain by using queries. The method of query selection to effectively define a research domain is beyond the scope of this paper and leaves an open room for future improvement. The last shortcoming of this approach is the existence of time lag. It takes time for a paper to be recognized by academia, and it also takes time from the completion of research to the publication. Without amendment from experts' opinions, the so-called current status may not be so current, and it can be difficult to grasp the current landscape correctly.

While the computer-based approach is a powerful tool to visualize the overall structure of target research domain in a way that even cannot be done by experts. The result obtained by this approach is nothing more than reference for those who contribute to the development of hydropower. We hope that our landscape can help promoting the research of hydropower by providing a fast and clear grasp of current status.

REFERENCES

[1] O. Paish, "Small hydro power: technology and current status," *Renew. Sustain. Energy Rev.*, vol. 6, no. 6, pp. 537–556, Dec. 2002.

[2] K. Börner, C. Chen, and K. W. Boyack, "Visualizing knowledge domains," *Annu. Rev. Inf. Sci. Technol.*, vol. 37, no. 1, pp. 179–255, 2003.

[3] K. W. Boyack, R. Klavans, and K. Börner, "Mapping the backbone of science," *Scientometrics*, vol. 64, no. 3, pp. 351–374, 2005.

[4] D. K. Okot, "Review of small hydropower technology," *Renew. Sustain. ENERGY Rev.*, vol. 26, pp. 515–520, 2013.

[5] J. W. Labadie, "Optimal operation of multireservoir systems: State-of-the-art review," *J. WATER Resour. Plan. Manag.*, vol. 130, no. 2, pp. 93–111, 2004.

[6] N. Kishor, R. P. Saini, and S. P. Singh, "A review on hydropower plant models and control," *Renew. Sustain. ENERGY Rev.*, vol. 11, no. 5, pp. 776–796, 2007.

[7] C.-J. Yang and R. B. Jackson, "Opportunities and barriers to pumped-hydro energy storage in the United States," *Renew. Sustain. Energy Rev.*, vol. 15, no. 1, pp. 839–844, Jan. 2011.

[8] H. Huang and Z. Yan, "Present situation and future prospect of hydropower in China," *Renew. Sustain. ENERGY Rev.*, vol. 13, no. 6–7, pp. 1652–1656, 2009.

[9] X. Li, H. Chen, Z. Huang, and M. C. Roco, “Patent citation network in nanotechnology (1976–2004),” *J. Nanoparticle Res.*, vol. 9, no. 3, pp. 337–352, Jan. 2007.
[10] M. Thelwall, K. Vann, and R. Fairclough, “Web issue analysis: An integrated water resource management case study,” *J. Am. ...*, vol. 57, no. 10, pp. 1303–1314, 2006.
[11] N. Shibata, Y. Kajikawa, and I. Sakata, “Extracting the commercialization gap between science and technology — Case study of a solar cell,” *Technol. Forecast. Soc. Change*, vol. 77, no. 7, pp. 1147–1155, Sep. 2010.
[12] M. E. J. Newman, “Fast algorithm for detecting community structure in networks,” *Phys. Rev. E*, vol. 69, no. 6, p. 66133, 2004.
[13] M. E. J. Newman and M. Girvan, “Finding and evaluating community structure in networks,” *Phys. Rev. E*, vol. 69, no. 2, p. 26113, 2004.
[14] G. Bruntland and others, “Our common future: The world commission on environment and development.” Oxford: Oxford University Press, 1987.
[15] R. E. N. 21 S. Committee and others, “RenewableS 2013: Global Status Report.” REN, 2013.
[16] M. Melikoglu, “Hydropower in Turkey: Analysis in the view of Vision 2023,” *Renew. Sustain. Energy Rev.*, vol. 25, pp. 503–510, Sep. 2013.
[17] Y. Kajikawa, J. Ohno, Y. Takeda, K. Matsushima, and H. Komiyama, “Creating an academic landscape of sustainability science: an analysis of the citation network,” *Sustain. Sci.*, vol. 2, no. 2, pp. 221–231, Jul. 2007.
[18] J. W. M. Rudd, R. Harris, C. A. Kelly, and R. E. Hecky, “Are hydroelectric reservoirs significant sources of greenhouse gases?,” *Ambio*, vol. 22, no. 4, pp. 246–248, 1993.
[19] U.S. Energy Information Administration, “International energy statistics, renewable electricity generation”, Retrieved 12/30/13, http://www.eia.gov/cfapps/ipdbproject/IEDIndex3.cfm?tid=2&pid=2&aid=12