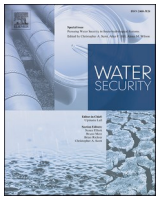
Water Security 16 (2022) 100117

Contents lists available at ScienceDirect

Water Security

journal homepage: www.sciencedirect.com/journal/water-security



Watery research boundaries: A bibliometric and network science approach to explore gaps and overlaps in water research

Lauren E. Dennis [a,b,1], Caitlin A. Grady [a,c,*,2]

[a] Department of Civil and Environmental Engineering, Pennsylvania State University, United States
[b] National Science Foundation LandscapeU Fellow, United States
[c] Rock Ethics Institute, Pennsylvania State University, State College, PA, United States

ARTICLE INFO

Keywords:
Bibliometric analysis
Networks
Water management
Transboundary water resources
Water security
Research collaboration

ABSTRACT

As the global community faces increasing water-related challenges, enabling safe and secure water access will require cooperation, purposeful water management, and a variety of expertise and interdisciplinary research collaboration. This review uses tools from bibliometrics and network science to explore overlap and collaboration of scholars studying transboundary water resources and water security management. We explore intersections between these communities through an analysis of publications trends, a content analysis of abstracts using natural language processing, and co-authorship networks. We glean five key findings from our results, including that slight variations in keywords used in the literature search in these two topic areas result in different communities of scholars and publications. Our results show that while publications on these topics are increasing over time and there is meaningful overlap between the two topics, the number of scholars publishing in both areas is not increasing over time. The co-authorship networks demonstrate that few authors participate in both transboundary water resources and water security management research communities, and that authors who have knowledge from both topic areas are uniquely positioned within their social networks to facilitate collaboration. We find no correlation between the betweenness centrality and the citation count for authors, measures which are both used to evaluate author influence. The content analysis of abstracts reveals important areas of overlap in the topics addressed, such as climate change, development, and governance, as well as areas of dissimilarity in the scales and focus of these works. Although we found that the broad scope of the water security framework included some of the most prominent scholars studying transboundary water resources, much of the transboundary water resource scholarship was not captured by water security keywords. This work demonstrates that if we are to continue to use integrative yet actionable frameworks in the pursuit of convergent water research, we must think carefully about how we craft these frameworks and whether our choice of language is constructive or destructive in bringing together relevant scholars and research.

1. Introduction

Under future climate change, an increasing proportion of the global population will face a variety of water-related challenges such as water scarcity and flooding [27]. Increased variability will necessitate skilled and purposeful water management, often requiring the cooperation of many parties, especially since about 40% of the global population lives in an international river basin shared among two or more countries [35]. As the study of water security transcends many scales [9,13] and is often transboundary in nature, and other similar themes such as the importance of good governance are often explored [9,13], we hypothesized that there would be substantial overlap and regular examples of collaboration between scholars who study water security management and scholars who study transboundary water resources. The need for interdisciplinarity and simultaneous consideration of diverse and interconnected issues in addressing global water challenges has been recognized by numerous scholars [10,33]. Previous studies have also shown that collaboration can increase the number of publications a researcher publishes [21] and that collaborative works tend to have a higher impact [18]. If we are to address grand water challenges, with

* Corresponding author at: 212 Sackett Building, The Pennsylvania State University, University Park, PA 16802-1408, United States.
E-mail address: cgrady@psu.edu (C.A. Grady).
[1] ORCID: 0000-0003-4384-4222.
[2] ORCID: 0000-0002-9151-6664.

https://doi.org/10.1016/j.wasec.2022.100117
Received 22 October 2021; Received in revised form 8 April 2022; Accepted 8 April 2022
Available online 15 April 2022

collaboration and convergent research aiding in this pursuit, we must determine how best to bring the relevant skills and expertise of many researchers together in a productive collaboration.

A first step in this process is to ensure diverse research communities can communicate effectively across language. Language has the potential to both constrain and facilitate collaboration, and the specific language used to describe various water-related concepts has evolved over time. A discussion at the Environmental Peacebuilding Conference held in Irvine, California, in October 2019 on the topic of definitions and language use inspired this work. The first high-level panel of this conference began by articulating how different parties defined Environmental Peacebuilding. Some panel members focused on environmental "security" as a framework and explicitly spoke about the environment and war, including preventative defense and environmental cooperation. Others were more interested in using the term to emphasize a relationship between the environment and healthy communities. One panel member claimed that a more generalizable definition should be created to unite the community, as people currently use the definition in the way that they like it best. While some panel members thought that coming to a shared agreement on the definition was important, others emphasized that the definition does not matter. Instead, in practice, they described that people need an "entry point" with which to realize that their work is relevant in multiple contexts.

In this study, we utilize tools from bibliometrics and network science to analyze overlap and collaboration between the study of transboundary water resources and water security management to examine whether there is a sufficient "entry point" for these two communities of researchers to come together. These two topic areas were frequently featured at the Environmental Peacebuilding Conference and seemed to explore similar themes, although we observed that methodological techniques or foci seemed to differ depending on whether the session focused on one topic or the other. We seek to expand understanding and stimulate discourse around how and why transboundary water resource and water security management research themes overlap and differ. First, we identify communities of researchers who study transboundary water resources and water security management through keyword searches in the Web of Science [8]. We then investigate overlap through examining the growth of publications in the two topic areas, exploring disciplinary scope and publication sources, examining the content of abstracts using natural language processing techniques to determine word frequency and word similarity, and constructing co-authorship networks and publication networks. Rather than providing a thematic analysis style review, our study expands current research across both transboundary water resources and water security management using these three techniques to explore the current state of collaboration and to showcase areas of intersection and dissimilarity that may provide potential opportunities for knowledge sharing.

2. Materials and methods

2.1. Database generation and categorical analysis

Search terms were selected to identify research within the Web of Science (WOS) database [8] related to water resources shared over boundaries. We included all publications that were returned from the keyword search when the search terms were present in the title, abstract, keywords, or "keywords plus" [7,29]. Our dataset was exported from WOS on May 13th, 2020 and analyzed using Python. Table S1 contains select phrases used in exploratory topic searches and the resulting number of publications. The two keyword searches ("Transboundary" & "Water Resources") and ("Water Security" & "Management") were chosen due to their similar size search results, and their inclusion of keywords comparing "transboundary" and "water security". Authors whose publications are found in the results of both topic searches in WOS are defined as "overlapping" authors in this study. Overlapping authors are examined as a proxy for research participation in each topic area of transboundary water resources and water security management. We examine author overlap by comparing the number of authors whose published literature appears in each topic search.

2.2. Content analysis

To examine overlap between topics, we analyze the content of abstracts for all publications found in the search results using natural language processing techniques from WordCloud [24] and Natural Language Toolkit (NLTK) [3] Python packages. First, we compare the 30 most commonly used words and phrases in each topic area. We use word frequency to examine potential overlap or differences between publications on transboundary water resources and water security, as we would expect that if the documents described similar content, their most frequently used words would also be similar. Next, we use the similar function within NLTK to examine the distributional similarity of important and relevant words, and we compare the similar words returned for each topic area. Distributional similarity measures the number of contexts shared by a word of choice and other words found within the body of text [26]. By examining other words found within common contexts as "transboundary," "security," and "management" in each topic area and comparing them, we are able to hypothesize about whether or not the words are being used in similar ways in each research community.

2.3. Network creation and analysis

In addition to categorical and content analyses, we utilized tools from network science, the study of complex phenomena through networks, to construct and analyze two different types of networks: co-authorship networks and a publication network (Fig. 1). Co-authorship networks were constructed with information gathered from each WOS search in which nodes represent authors and edges between authors indicate co-authorship (Fig. 1). The average node degree (number of edges) represents the average number of co-authors or research collaborations, edge weights represent the frequency of collaborations, the number of connected components represents the number of isolated co-author communities or singular authors, and we use the size of the largest connected component as an indication of how connected researchers are to the rest of the field [11,25]. We examine the network structure of each topic separately as well as a combined network in which authors that appear in both topic searches are denoted. We also examine network position by computing the betweenness centrality of nodes as a potential proxy for control over communication activity within the network [1]. Betweenness centrality is a measure of the frequency in which a node appears in the shortest paths of other nodes [12]. Building on Chung's (2009) interpretation of Burt's (1992) structural holes theory in which "power and influence accrue to those who broker connections between unconnected groups of people" [6], we examine betweenness centrality as it relates to a node's topic categorization and connection to others who publish in a particular topic.

Betweenness centrality may also represent a measure of structural social capital held by a researcher within the network [23]. Li et al. [23] describe relational social capital as assets that are shared through relationships, which can be manifested in trust, commitment, and reciprocity, and expressed as repeated co-authorships [23]. In our network, this is expressed as the edge weight. Cognitive social capital can be expressed in a co-authorship network through the extent of time that an author has published in a particular discipline [23]. After an author spends time in a discipline, they gain knowledge of "shared visions, traditions, codes, languages, knowledge, interpretations, systems of meanings, social networks, and collectively owned social capital" [23], which they can then share with others. Because we assign authors to explicit research communities through the bounds of our topic searches within the Web of Science publication database, we can measure cognitive capital through both publication tenure and topic expertise.

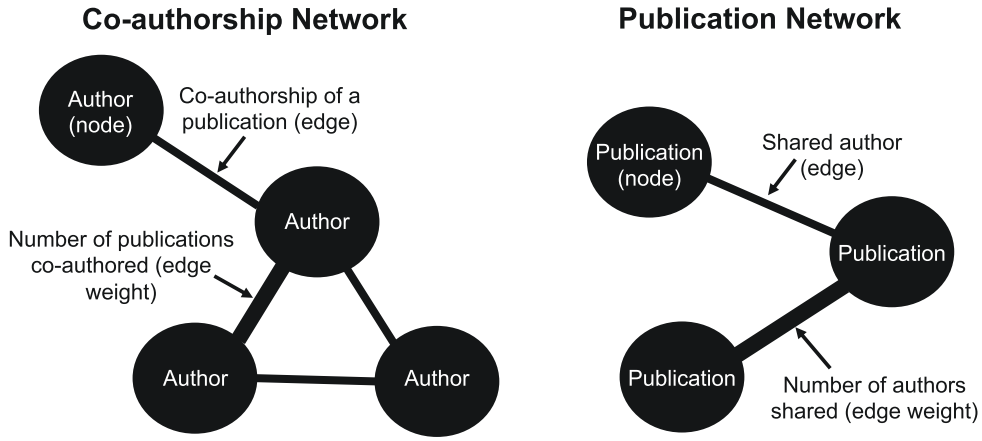


Fig. 1. A schematic diagram of both types of networks analyzed in this work. In the co-authorship network, nodes represent authors, and edges between authors signify co-authorship of a publication. In the publication network, nodes represent published works and nodes share an edge if they share an author.

We also calculate the total number of citations each author receives for their publications in either topic and examine whether this correlates with an author's betweenness centrality. If we use citation count as a potential measure of influence [17], we would expect that it would be correlated with betweenness centrality, as this is a measure of potential structural influence within the network.

We also construct and analyze a network in which publications are nodes, edges are instances of co-authorship among one or more authors in the publication, and node color represents the topic area (Fig. 1). Here, we seek to understand how particular ideas or frameworks for understanding water resources are connected by research collaborations. We examine how publications from each topic area are linked to one another through authors involved in multiple topic areas to understand the potential for information flow. All networks were created using the NetworkX Python package [16] and visualized using Gephi [2]. We utilized the Fruchterman-Reingold force-directed algorithm within Gephi to create network layouts. Code to create the networks from WOS search results can be found in the SI.

3. Findings and discussion

Our results lead us to consider 5 main findings outlined below and featured in Fig. 2.

3.1. Content and author overlap

Studies of transboundary water resources (TWR) and water security and management (WSM) have meaningful overlap, even though few authors publish in both topics. Analysis of results from the Web of Science searches reveals overlap in the general disciplinary reach for each topic among the four most prevalent disciplinary categories (Fig. 3a). A wide variety of disciplines study each topic, but the distribution of categories is larger for the WSM topic area. TWR publications fall into 65 different WOS categories, whereas WSM publications fall into 103

Finding	Process	Implication
3.1 Studies of TWR and WSM have **meaningful overlap**, even though few authors publish in both topics.	Content and categorical analysis	Both topics explore themes such as climate change, development, and governance, while the scales and focus on types of water infrastructure differ, suggesting that there are opportunities for knowledge sharing.
3.2 Even slight variations in the **keywords** used in a literature search in these topic areas produce different communities of scholars and publications.	Categorical analysis	The use of field-specific language could make this research less discoverable, and because TWR and WSM research is found in many different journals, there may be limited opportunities for organic interaction.
3.3 Analysis of co-authorship networks show that **collaborations** vary in scale and frequency, and WSM may have a broader scope which includes some of the most prominent scholars in TWR.	Co-authorship networks and categorical analysis	While WSM may be a broader framework and collaborations among water security scholars tend to be larger, many TWR scholars do not appear in WSM literature, and there may be lessons learned from TWR scholars about working across boundaries that could benefit WSM scholars.
3.4 Authors who publish in **both topic areas** have a **higher betweenness centrality**, although these scholars may not be seen as traditionally influential based on their citation count.	Co-authorship networks and categorical analysis	Scholars who are active in more than one topic area may be able to facilitate collaboration and shared language use utilizing cognitive capital acquired in other disciplines, although these scholars may not be viewed by the community as the most traditionally influential.
3.5 Outside factors influence the organization of some clusters of co-authors within the network in which all authors share an institutional affiliation or research funding source.	Co-authorship networks	There may be particular institutions or research funding structures that are uniquely positioned to further convergent water research.

Fig. 2. Summary table of main findings, the methods employed which support the finding, and the implications of the finding.

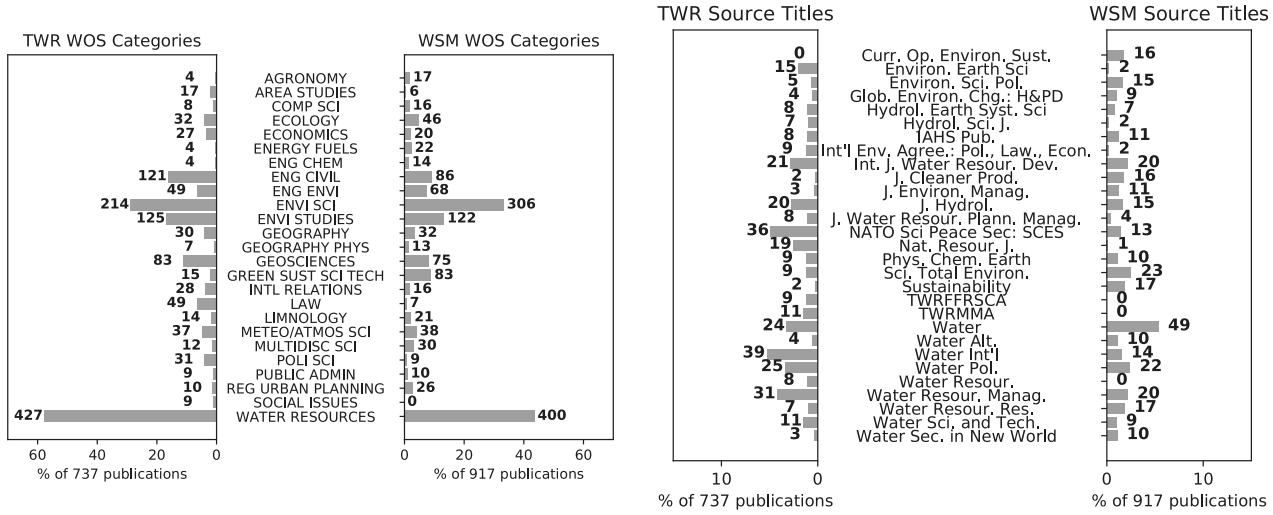


Fig. 3. WOS Categories and Source Titles for Literature in "Transboundary" & "Water Resources" or "Water Security" & "Management". a) The top 20 WOS disciplinary categories for publications in the TWR topic search or WSM topic search are shown, where bar length indicates the percentage in that category out of all publications in a topic area, and bar labels are publication counts within that category. b) The top 20 source titles for TWR and WSM topic searches are shown as a percentage of total publications (bar length), and the number of publications is shown with the bar label. Journal abbreviations can be found in the Supporting Information.

different categories, and TWR only has 7 unique categories. While 58 categories overlap, some notable differences include 9 publications that fall under "Social Issues" within TWR literature compared with zero WSM publications in this category. Similarly, 6.7% of publications from TWR compared to only 0.8% of publications from WSM fall into the category of "Law." Although addressing water security can most certainly involve social issues and law (e.g. Kuokkanen, 2017 [38]), it appears that this is not captured within the WOS categorization. Overall,

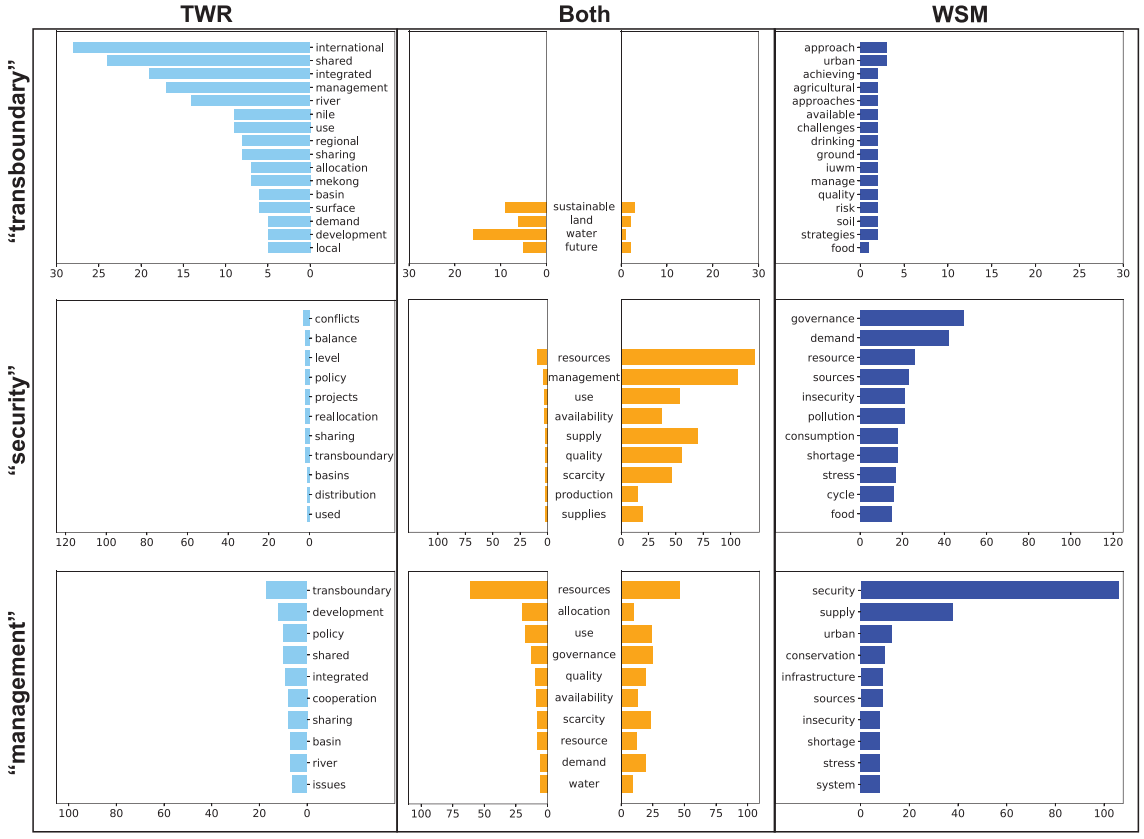


Fig. 4. The top 20 words that appear in the same context as "transboundary," "security," and "management," as reported by NLTK's similar function. Each row represents the word whose similarity is examined, and each column shows the topic of the abstracts examined. The x-axis shows the number of contexts shared by the word labeled on the y-axis and the input word. "iuwm" (found in the top left plot) is an abbreviation for integrated urban water management, although the measure of similarity refers to the abbreviation and not the extended phrase.

the overlap between the top 4 most frequent categories between the two topic searches, and the general similarity between categorizations, indicate an apparent similarity between the disciplinary reach of the two topic searches.

Exploring the language used in abstracts in each topic area allowed us to better understand content similarities and differences between the two topic areas, and we found many areas of overlap as well as some important distinctions relating to scale and focus. Among the most frequently used 100 words and phrases within abstract text in "Transboundary" & "Water Resources" and "Water Security" & "Management" search results, a total of 68 words/phrases are shared (Table S3). Our results suggest that both climate change and development are particularly salient themes within TWR and WSM scholarship, as these words have high frequency across both groups of articles (Table S3).

Beyond word frequency, an examination of words that appear in the same contexts as specific keywords showed us similarities and differences in the ways in which these words are used in each body of abstracts (Fig. 4). When comparing words which share the same contexts as the word "transboundary" in each collection of abstracts, "sustainable," "land," "water," and "future" are found in both collections (Fig. 4). Contexts relating to scale differ when referencing "transboundary" in each group: TWR features words such as "international," "regional," "Mekong," "basin," and "local," whereas "urban" is the only word similar to "transboundary" in WSM abstracts referring to scale. "IUWM" also appears, referring to integrated urban water management, again at the urban scale. Words associated with the context of water sources differ between collections of abstracts: words such as "river" and "surface" share contexts with "transboundary" in TWR abstracts, whereas words such as "ground" and "drinking" share contexts with "transboundary" in WSM abstracts. We interpret these results to mean that WSM literature is more often focused on smaller scales, whereas TWR literature focuses on state-to-state scales and larger hydrologic features such as rivers or large, international basins. This supports findings from previous reviews which found that water security research at transboundary scales is significantly less frequent than other scales [9,13]. Cook and Bakker [9] highlighted that scholars within different disciplines tend to highlight different scales; studies of development most often focused on the national scale, hydrological studies occurred at the watershed, regional, or national scale, and social science studies were most often focused at the community scale [9]. Work which examines conflict and cooperation over transboundary water resources often highlights the state-to-state scale, largely because this work deals with water resources in the complex context of institutional, political, social, and cultural systems, although some of this work also describes sub-national scales [31,34].

There are 10 words that are found in similar contexts to the word "management" that appear in both TWR and WSM abstracts (Fig. 4). Many of these words are associated with the quality and quantity of water, and words relating to human interaction with the water system appear such as "governance." WSM abstracts return more words with a negative connotation with similarity to "management" such as "shortage," "insecurity," and "stress," whereas these words are not found in the 20 most similar words in TWR abstracts. TWR abstracts contain words similar to "management" that are related to human negotiation over water resources such as "shared," "sharing," or "cooperation," while WSM abstracts contain words that seem to refer more to the water system itself such as "infrastructure" and "conservation." Our results also suggest that the inclusion of the keyword "management" with "water security" tends to highlight research at urban scales (Fig. 4). We interpret our similarity results around the context of "management" to show a difference between topics, where WSM abstracts focus more on how the water infrastructure is managed, and TWR literature focuses more on the social nature and negotiation taking place within water management such as studies on conflict and cooperation over water resources [34]. Definitions of water security provided by entities such as the Global Water Partnership and the UN recognize that management and negotiation are key components of water security [15,32], and previous water security scholarship has called for a more "integrative and broad framing" of water security research, while recognizing the difficulties in operationalizing this broad framing and the prevalence of "reductionist" studies [9,36]. Cook and Bakker describe that an integrative approach to water security highlights the topic of governance, advancing water security research [9], while governance, coordination, and collaboration have long been main themes within research on transboundary water resources [34]. Wolf describes that increasing threats to water resources related to water quality and the use of less traditional sources that are not constrained to watershed boundaries could distinguish future water disputes from those in the past [34], which could more tightly link scholarship on transboundary water resources and water security.

In exploring whether the words "transboundary" and "security" were being used in the same way in TWR and WSM literature, we noticed that these words shared many contexts with words representing other systems such as "food" and "agricultural" within the WSM literature. In TWR abstracts, however, "transboundary" and "security" shared more common contexts with words representing the water system itself, such as "basin" or "river." These results suggest that the conceptualization of security of the water system may extend to include other systems explicitly within water security literature, whereas other systems are mentioned less often or explicitly within TWR literature. Previous work has highlighted water security in the context of the water-energy-food (WEF) nexus, suggesting that water security is often paramount to ensure food and energy security [30].

It is likely that the conceptualization of research on TWR and WSM will continue to evolve, as both TWR and WSM appear to be relatively new terms in the literature. TWR emerges within two publications in 1991 and WSM emerges from three publications in 1995 (Fig. 5). The first instance in which an author published works in both topic areas occurs in 2000 (Fig. 5). A total of 4,577 authors published in either topic area, with 1,551 publishing only on the topic of TWR, 2,832 only publishing on the topic of WSM, and 194, representing 4.2% of the total, publishing in both topics between 1991 and 2019. The percentage of authors publishing in both topic areas in a given year ranges from 0.8 to 6.2% from 2000 to 2019. The percentage of authors publishing in both topic areas does not appear to be increasing from 2015 to 2019, and between 1991 and 2007, only 7 authors published in both topic areas. Thus, while the disciplinary reach and similar content between abstracts suggests that there is substantial overlap between the topics, few

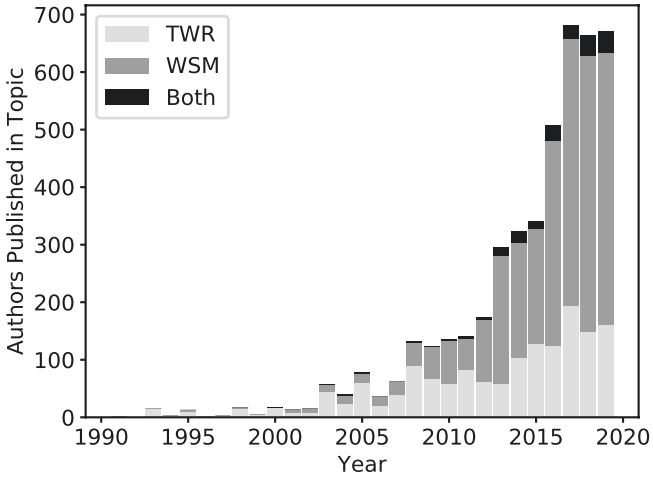


Fig. 5. The number of authors that have published in the ("Transboundary" & "Water Resources") topic search or the ("Water Security" & "Management") topic search in a particular year. Authors belong to the "Both" category if they have a publication in each topic area, although their appearance in a particular year may occur from a publication in one topic area or another.

scholars are publishing in both areas.

3.2. Effect of variations in keywords

One potential reason that scholars would not appear in search results for both TWR and WSM, even if their work might fall under these topic areas, is that they are not using the same terminology or keywords to describe their work. Our results suggest that even slight variations in the keywords used in a literature search in these topics produce different communities of scholars and publications (Table 1, Table S1). As shown in Table 1, very few authors overlap between topic searches in general, even between searches in a similar topic that use seemingly similar language. Comparing searches that each use the word "governance" produces some author overlap: 969 authors overlap between the "Water Resources" & "Governance" search and the "River Basin" & "Governance" search (30% of the authors in the "Water Resources" & "Governance" search and 45% of the authors in the "River Basin" & "Governance" search). When comparing keywords used for searches that address the two main topic areas considered in this paper, we find very little overlap. As mentioned previously, only 194 authors overlap between "Transboundary" & "Water Resources" and "Water Security" & "Management" searches (11% and 6% of authors in each search). Previous work has shown that author chosen keywords are likely to include field-specific language [39]. One study examining the field specificity of keywords in 200 articles found that the percentage of keywords described as "general terms" was very low (only 5.7% of words used in titles and only 21% of keywords) [37]. Using field-specific language that is not well understood by many disciplines could prevent the sharing of knowledge among relevant communities, and especially in a web-based search of literature, could make research less discoverable.

In addition to field-specific language, the language used to describe broad research themes changes over time. Although the first instances of research described by "Transboundary" & "Water Resources" and "Water Security" & "Management" keywords appear in 1991 and 1995 respectively in Web of Science, the first search result for "River Basin" & "Management" appears in 1963. This could indicate that research on the themes described in this work has occurred long before 1991, but the language has changed or expanded. This makes language choice very important and also provides grounds for an investigation into whether this language could be a barrier in uniting communities of scholars and instigating collaboration on important research addressing grand water challenges. Authors who publish in multiple topic areas addressing both transboundary water resources and water security management, who may be familiar with the language spoken in both communities, have the potential to share ideas and knowledge gained from both topic areas with other researchers. This furthers convergent water research, and a lack of author overlap between topic searches might suggest fewer opportunities for knowledge sharing within these communities (Table 1).

An additional barrier to knowledge sharing between TWR and WSM could be that scholars in each topic area are publishing their work in different sources. Publication sources returned from each topic search share four of the top 10 most popular journals (Fig. 3b). Overall, 128 publication sources feature publications from both TWR and WSM topic searches. As showcased by the fact that not a single source has published more than 49 publications on either topic, there is a large distribution of publication sources for both topic searches. WSM literature returns a greater number of different publication sources (461 sources), while TWR literature covers 364 different outlets. A lack of overlap may suggest that ideas shared on each topic may not reach scholars in the other topic if they do not frequently read publications from the same sources; thus, potential for knowledge sharing may be greater in sources that contain publications from both topics.

Table 1
Author Overlap by Topic Area.

Subject	1)	2)	3)	4)	5)	6)
1) "Transboundary" & "Water Resources"	1,745	375	217	252	194	141
2) "Water Resources" & "Governance"		3,272	969	503	433	412
3) "River Basin" & "Governance"			2,130	701	253	214
4) "River Basin Management"				3,894	254	157
5) "Water Security" & "Management"					3,026	1,102
6) ("Water Security" & "Governance") OR ("Water Security" & "Policy")						1,752

Note. Topic searches are numbered and identical between the horizontal and vertical axis. Each location within the matrix represents the number of authors which appear in both topic searches indicated by the corresponding axes. For example, 217 authors appear in in both the 1) "Transboundary" & "Water Resources" search and the 3) "River Basin" & "Governance" topic search ($a_{1,3}$). The main diagonal represents the total number of authors in that particular topic search.

3.3. Variation in collaboration scale and scope

Analysis of co-authorship networks show that collaborations vary in scale and frequency, and WSM may have a broader scope which includes some of the most prominent scholars in TWR. Our water security and management co-authorship network had 3,026 authors, while the transboundary water resources co-authorship network contained 1,745 authors (Table 2, Fig. 6). The WSM network had a higher average number of authors per paper, and the average node degree is also greater in the WSM network. The greater average node degree suggests that co-authorship is more frequent in the WSM network than the TWR network, since average node degree can be a proxy for research collaboration [25]. Nodes representing authors who publish in both topic areas have a higher average node degree (11.6 unweighted, 12.5 weighted) which could indicate that authors who publish in both topic areas are more collaborative, although the structure of our co-authorship network naturally includes more collaborators for authors who publish in both topic areas. The greatest number of citations for an author who has published works in the WSM network is 2,809, whereas the most highly cited author who has published works in the TWR network has 378. The most highly cited author in the TWR network publishes in both topic areas. In contrast, the most highly cited author in the WSM network only published in the area of WSM. From our analysis of the co-authorship networks statistics, we conclude that the WSM network appears to have larger collaborations than the TWR network, the WSM network contains a few very highly cited authors who do not publish in TWR, whereas the most highly cited author in the TWR network also publishes within WSM. We interpret these results to mean that WSM could be a broader topic area under which some TWR literature coincides.

Table 2
Co-authorship Network Statistics.

Statistic	TWR	WSM	Combined Network
Number of nodes	1,745	3,026	4,577 (194 nodes appear in both topics)
Number of edges	4,066	11,493	15,405
Average node degree (weighted/unweighted)	4.94/4.66	7.81/7.6	6.95/6.73
Number of connected components	457	534	882
Size of the largest connected component	279 nodes	761 nodes	1,409 nodes

Note. Network measures computed for co-authorship networks on the topic of "Transboundary" & "Water Resources" (TWR), "Water Security" & "Management" (WSM), and a combined network in which authors from each topic search are included. Single authors are included in these networks as singular nodes, and in the calculation of the number of connected components, a single node is considered one component.

3.4. *Influence of authors who publish in both topic areas*

Authors who publish in both topics have a higher normalized, average betweenness centrality than those who fall in only one topic area by two orders of magnitude (0.001 versus 4×10^{-5} for TWR and 8×10^{-5} for WSM). However, these authors may not be seen as traditionally influential based on their citation count. Fig. 6 (panel a.) showcases instances within the network where a node in the "both" category serves as a bridge between two topic clusters which would otherwise be disconnected. The existence of authors that connect the two topic areas suggests that the two topic areas are likely meaningfully connected or relevant to one another somehow, since authors tend to publish works within a general scholarly domain.

The publication network emphasizes the key role that authors who publish in both topic areas could play in connecting researchers and relevant bodies of work: 24% of the authors in the publication network fall into the "both" category and link papers of different topic areas, whereas these authors only make up 4% in the co-authorship network (Fig. 6 panel a.). There are only 19 papers that are returned in both topic searches; thus, only 19 nodes in this publication network fall into the "both" category. Many of these nodes are linked to publications in either topic, demonstrating that the authors who published that work are active in other areas and came together to collaborate (Fig. S4 panel b.). Interestingly, some nodes which appear in both topic searches are singular nodes within the publication network, indicating that the authors of that publication have not published other work on the topics of transboundary water resources or water security and management beyond that one publication which covers both topics.

Because betweenness centrality can be seen as a proxy of power and influence as a controller of information [1,6], we also investigated the relationship between betweenness centrality and author citation count within publications returned by the topic search. While the author with the highest citations has the second-highest betweenness centrality in the TWR network, the author with the highest citation count in the WSM network has the 34th highest betweenness centrality. We observe no linear relationship between the betweenness centrality and citation count for each node ($R^2 = 0.022$, Fig. S5). The lack of relationship suggests that some nodes which are influential based on citations are not influential based on betweenness centrality, and vice versa.

If we view the citation count of an author as a potential measure of influence [17], we would expect to see betweenness centrality and citation count correlated in these data. Our results challenge this previous work since we found that betweenness centrality and citation count are not correlated within our network. Previous scholarship suggests that outside factors influence citation practices, and in the context of our network, we believe that this could explain why we did not find

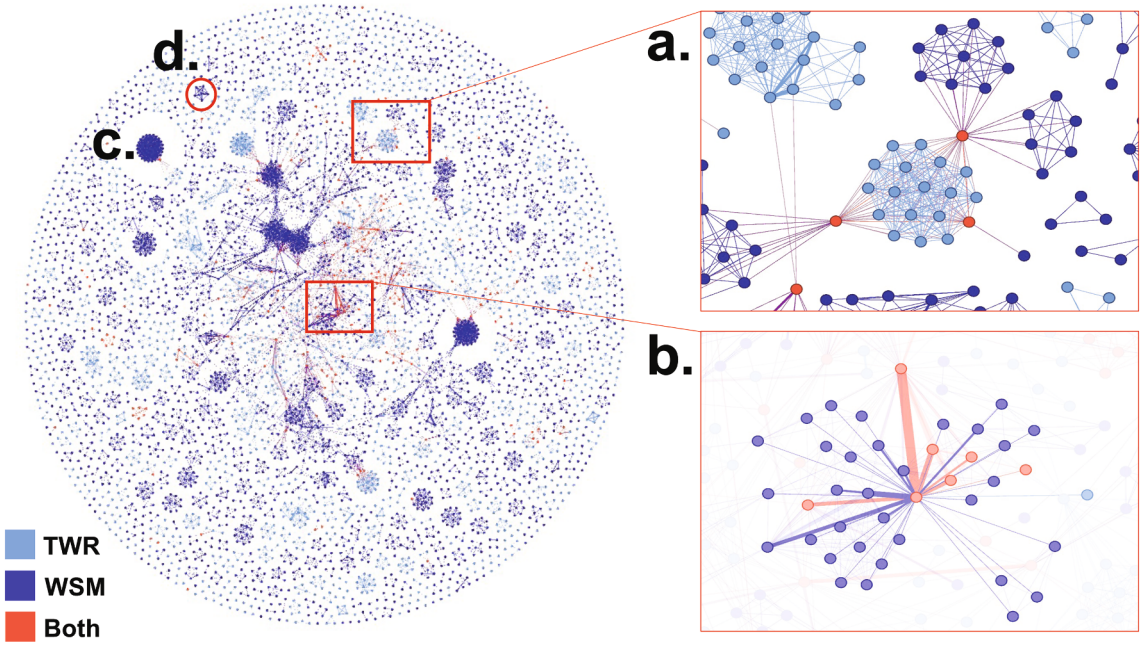


Fig. 6. The structure of research collaboration within topics of TWR and WSM as depicted by a combined co-authorship network. Nodes are authors, edges indicate co-authorship, and the node color indicates the topic search in which that author appeared. Edge thickness indicates the frequency of collaboration between authors. Panel a. shows an example of authors who appear in both topic areas serving as a bridge between different topic areas, and panel b. shows an example of a node in the both category whose neighbors that fall in the both category share the same institutional affiliation (except for one). Location c. shows an instance where 53 authors collaborated on a single paper, and location d. shows repeated co-authorship from the same research funding source.

correlation between betweenness centrality and citation count. The prestige of an institution [22] or journal may have influence over the number of citations a publication receives [28], and it has been shown that publications from authors from traditionally underrepresented groups may receive a disproportionately low number of citations. For example, women tend to be cited less than their male counterparts 1) in countries with high research productivity even when women are in the same roles [20], 2) in Astronomy even when paper properties are controlled for [4], and 3) in the field of engineering and technology, even when articles by women engineers were published in journals with higher impact factors [14]. Within the field of communications, Chakravartty et al. [5] found that non-white authors were cited significantly fewer times than their white counterparts [5]. Variation in self-citation practices among groups may also influence citation count: for example, men tend to cite themselves more frequently than women [19]. Thus, although citation count may be used as a measure of influence in some cases, it is an imperfect measure that does not always reflect the power and influence of an author. We emphasize that in our network, citation count does not accurately reflect the power and influence of authors with unique capability to further convergent research. The factors mentioned above, among others, may explain a lack of correlation between citation count and betweenness centrality.

3.5. Institutional factors influencing particular co-author clusters

Outside factors influence the organization of some clusters of co-authors within the network in which all authors share an institutional affiliation or research funding source. The network structure includes a handful of very dense, large clusters of nodes (Fig. 6). Many of these clusters represent a large research collaboration that resulted in a single paper. One example of this is depicted in Fig. 6 (location c.), in which 53 authors collaborate on a single paper, and only one of the authors in that cluster has published additional work within either topic area depicted in the network (in this case, in the other topic area). We also identified instances in which repeated collaborations could be traced to multiple publications from the same research funding source (Fig. 6 location d.), suggesting that sometimes research funding may influence research collaboration frequency or structure. Overall, however, the collaboration frequency is weak. The average collaboration frequency represented by the average edge weight is one, and the distribution of edge weights is skewed toward a single collaboration instance (Fig. S2).

In some instances, we see that institutions influence the organization of nodes that fall in the "both" category. The centermost node in Fig. 6 (panel b.) shows an example of an author who publishes in both topics whose neighbors also publish in both topics. Each of these authors share an institutional affiliation except for one. We found multiple examples of this, suggesting that perhaps particular institutions are more likely to have scholars who publish in both topic areas and collaborate. Although we could not discern why using our methods, we surmise that this trend may occur for a variety of reasons: 1) these institutions may recruit scholars whose research is cross-cutting and the institution may have a legacy of scholarship in these topic areas, 2) co-authorship practices at the institution may be such that co-authorship is frequently awarded to colleagues at the institution, or 3) there is a particular culture of collaboration at the institution that supports frequent co-authorship with colleagues at the institution. The first and third hypotheses may suggest that institutions could play a role in facilitating collaboration that could aid in research progress in important, convergent, global water-related challenges under future and current climate change. These institutions may have a focus that captures broader water-related work, which is well suited to the similarities in scope between the two topic areas we include in this study. This highlights an interesting premise for future research to elucidate ways in which institutions can influence such convergent approaches.

4. Future work and limitations

We have provided a precursory review of overlap and dissimilarities between TWR and WSM scholarship utilizing tools from natural language processing and network science, but we mainly highlight the current state of research collaboration. Future research has the potential to unlock further understanding in this space. We highlighted authors who publish in each domain as potentially important facilitators of collaboration and brokers of knowledge, but we wonder, does this facilitation and "translation" of field specific language actually occur? Does the work written by scholars who publish in both topic areas advance water research in a way that the work written by scholars who only publish in one topic does not? Does the involvement of an "interdisciplinary" scholar add richness or innovation to the scholarship? In our study, we assume that co-authorship implies collaboration, though this does not necessarily measure the varying degrees to which co-authors work together in practice. Further analysis of the publications included in the dataset in this study may reveal whether there is a real benefit to the scholarship resulting from co-authorship and meaningful collaboration between authors in different topic areas. Additionally, scholars in these areas could opine on further frameworks, tools, or knowledge that could be particularly useful in being shared between these communities. To advance future work, we have included the dataset and code inclusive of our analysis (SI, https://doi.org/10.26207/6qpj-e819).

The methods employed in this work are likely useful to explore overlap between other scholarly communities, as we chose two communities represented by two sets of search terms as an illustrative example but could have chosen any number of combinations of search terms. We selected TWR and WSM as keywords based on inspiration from the Environmental Peacebuilding Conference, yet the methods we developed could be replicated to leverage helpful exploration across and between any fields. We hope that this work will stimulate discourse about "entry points" within research frameworks and communities and inspire future work which explores more terminology. Multiple scholars have shown that identifying relevant bodies of literature based on search terms can be challenging. Issues such as keyword choice for particular paradigms evolving over time (Bentley, 2008), variance in author keyword choice (Kipp, 2006), excessive specificity of keywords (Babaii & Taase, 2013; Heckner et al., 2008), and limitations of databases (Müngen & Kaya, 2018) can make it difficult for scholars from one community to find relevant literature from another. This reflects both a need for careful thought by authors in their choice of keywords in publications, as discussed earlier in this work, and also a limitation in our work in that the scholarly communities represented are based upon search terms and are likely incomplete.

The scope of this study is limited to academic literature found in WOS. We acknowledge that full research participation in a topic area may not be captured by publication results in WOS from a keyword search, however, we argue that this is still a relevant analysis because scientific knowledge is often gained and shared through the published literature. We also acknowledge that "grey" literature contributes significantly to the topics featured in this article, and that the important influence of this body of work is not captured in our study.

5. Conclusions

From our five findings, we would like to leave the readers with three main conclusions. First, the use of language in scholarly publications can have important implications in how knowledge is shared, especially within water resources topics. Even a slight alteration in the keywords used in a search in a scholarly database causes a large variance in the publication results returned (Table 1). If a scholar is not familiar with the language or keywords used in another sub-discipline or field, they may miss out on what would otherwise be relevant work and knowledge from other topic areas or disciplines. This suggests that researchers must

be purposeful in the keywords they use, and that utilizing generalized keywords may facilitate the spread of knowledge to a broader but still relevant audience. Second, scholars who are active in more than one topic area or discipline may be able to facilitate collaboration utilizing cognitive capital that they have acquired in each discipline to facilitate the use of a shared language. This is not to say that all scholars must be interdisciplinary, but that these scholars have a unique role to play in executing convergent research. In our co-authorship and publication networks, we find that authors who publish on topics of both transboundary water resources and water security have a unique role in connecting what would otherwise be disparate scholarly communities on either topic. While these authors hold a high level of influence in the collaboration network as measured by betweenness centrality, this measure of influence is not correlated with a traditional measure of influence through citation count. Finally, our analysis of the content of abstracts reveals relevant intersections between topics of transboundary water resources and water security management in areas such as climate change, development, sustainability, management, governance, water availability, and other topics. Differences in foci between the two topic areas included the scale, systems, and the type of water infrastructure considered. Although we found that the broad scope of the WSM framework included some of the most prominent scholars in TWR, much of the TWR scholarship was not captured by these keywords (Table 1). If we are to continue to use integrative yet actionable frameworks in the pursuit of convergent water research, we must think carefully about how we craft these frameworks and whether our choice of language is constructive or destructive in bringing together relevant scholars and research.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors would like to thank the three anonymous reviewers of this work for their careful and constructive feedback. In fact, the review process for this article was the most helpful and constructive that either author has ever experienced. This article was significantly strengthened by the review process and we are very appreciative of the reviewers' time. The authors would also like to thank Sarah Rajtmajer for her technical assistance with initial data processing and network creation.

This work was partially supported by the National Science Foundation Award #1828822.

Datasets for this research are available at doi:10.26207/6qpj-e819. Code used for analysis is available in the Supporting Information.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.wasec.2022.100117.

References

[1] A. Abbasi, J. Altmann, On the Correlation between Research Performance and Social Network Analysis Measures Applied to Research Collaboration Networks, in: *2011 44th Hawaii International Conference on System Sciences*, 2011, pp. 1–10, https://doi.org/10.1109/HICSS.2011.325.

[2] M. Bastian, S. Heymann, M. Jacomy, Gephi: An Open Source Software for Exploring and Manipulating Networks, Retrieved from. International AAAI Conference on Web and Social Media; Third International AAAI Conference on Weblogs and Social Media, 2009.

[3] S. Bird, E. Klein, E. Loper (2009). *Natural Language Processing with Python.* Sebastopol: O'Reilly Media, Inc. Retrieved from http://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=443090.

[4] N. Caplar, S. Tacchella, S. Birrer, Quantitative evaluation of gender bias in astronomical publications from citation counts, Nat. Astron 1 (6) (2017) 1–5, https://doi.org/10.1038/s41550-017-0141.

[5] P. Chakravartty, R. Kuo, V. Grubbs, C. McIlwain, #CommunicationSoWhite, J. Commun. 68 (2) (2018) 254–266, https://doi.org/10.1093/joc/jqy003.

[6] K.S.K. Chung, L. Hossain, Measuring performance of knowledge-intensive workgroups through social networks, Proj. Manag. J. 40 (2) (2009) 34–58, https://doi.org/10.1002/pmj.20115.

[7] Clarivate. (2020a). KeyWords Plus generation, creation, and changes. Retrieved February 7, 2020, from https://support.clarivate.com/ScientificandAcademicResearch/s/article/KeyWords-Plus-generation-creation-and-changes?language=en_US.

[8] Clarivate. (2020b). Web of Science. Retrieved November 14, 2019, from https://clarivate-com.ezaccess.libraries.psu.edu/webofsciencegroup/solutions/web-of-science/.

[9] C. Cook, K. Bakker, Water security: Debating an emerging paradigm, Global Environ. Change 22 (1) (2012) 94–102, https://doi.org/10.1016/j.gloenvcha.2011.10.011.

[10] W.J. Cosgrove, D.P. Loucks, Water management: Current and future challenges and research directions: Water management research challenges, Water Resour. Res. 51 (6) (2015) 4823–4839, https://doi.org/10.1002/2014WR016869.

[11] C.K. Fatt, E.A. Ujum, K. Ratnavelu, The structure of collaboration in the Journal of Finance, Scientometrics 85 (3) (2010) 849–860, https://doi.org/10.1007/s11192-010-0254-0.

[12] L.C. Freeman, A set of measures of centrality based on betweenness, Sociometry 40 (1) (1977) 35, https://doi.org/10.2307/3033543.

[13] A.K. Gerlak, L. House-Peters, R.G. Varady, T. Albrecht, A. Zúñiga-Terán, R.R. de Grenade, C. Cook, C.A. Scott, Water security: A review of place-based research, Environ. Sci. Policy 82 (2018) 79–89.

[14] G. Ghiasi, V. Larivière, C.R. Sugimoto, W. Glanzel, On the Compliance of Women Engineers with a Gendered Scientific System, PLoS ONE 10 (12) (2015) e0145931.

[15] Global Water Partnership. (2000). *Towards water security: a framework for action.* (World Water Forum, Ed.). Stockholm: GWP.

[16] A.A. Hagberg, D.A. Schult, P.J. Swart, (2008). Exploring Network Structure, Dynamics, and Function using NetworkX. In G. Varoquaux, T. Vaught, & J. Millman (Eds.), *Proceedings of the 7th Python in Science Conference* (pp. 11–15). Pasadena, CA USA.

[17] J.E. Hirsch, An index to quantify an individual's scientific research output, Proc. Natl. Acad. Sci. 102 (46) (2005) 16569–16572, https://doi.org/10.1073/pnas.0507655102.

[18] J.S. Katz, D. Hicks, How much is a collaboration worth? A calibrated bibliometric model, *Scientometrics* 40 (3) (1997) 541–554, https://doi.org/10.1007/BF02459299.

[19] M.M. King, C.T. Bergstrom, S.J. Correll, J. Jacquet, J.D. West, (2016). Men Set Their Own Cites High: Gender and Self-citation across Fields and over Time. https://doi.org/10.1177/2378023117738903.

[20] V. Larivière, C. Ni, Y. Gingras, B. Cronin, C.R. Sugimoto, Bibliometrics: Global gender disparities in science, Nature News 504 (7479) (2013) 211, https://doi.org/10.1038/504211a.

[21] S. Lee, B. Bozeman, The Impact of Research Collaboration on Scientific Productivity, Soc. Stud. Sci. 35 (5) (2005) 673–702, https://doi.org/10.1177/0306312705052359.

[22] R. Leimu, J. Koricheva, What determines the citation frequency of ecological papers? Trends Ecol. Evol. 20 (1) (2005) 28–32, https://doi.org/10.1016/j.tree.2004.10.010.

[23] E.Y. Li, C.H. Liao, H.R. Yen, Co-authorship networks and research impact: A social capital perspective, Res. Policy 42 (9) (2013) 1515–1530, https://doi.org/10.1016/j.respol.2013.06.012.

[24] A. Mueller, (2020). WordCloud for Python documentation — wordcloud 1.8.1 documentation. Retrieved March 31, 2021, from https://amueller.github.io/word_cloud/.

[25] M.E.J. Newman, Scientific collaboration networks. I. Network construction and fundamental results, Phys. Rev. E 64 (1) (2001), 016131, https://doi.org/10.1103/PhysRevE.64.016131.

[26] nltk.text — NLTK 3.5 documentation. (2020). Retrieved December 18, 2020, from https://www.nltk.org/_modules/nltk/text.html.

[27] R.K. Pachauri, L. Mayer, Intergovernmental Panel on Climate Change (Eds.). (2015). *Climate change 2014: synthesis report*. Geneva, Switzerland: Intergovernmental Panel on Climate Change.

[28] H.P.F. Peters, A.F.J. van Raan, On determinants of citation scores: A case study in chemical engineering, J. Am. Soc. Inf. Sci. 45 (1) (1994) 39–49, https://doi.org/10.1002/(SICI)1097-4571(199401)45:1<39::AID-ASI5>3.0.CO;2-Q.

[29] M. Ruccolo, (2020). LibGuides: Web of Science Core Collection: Search Tips. Retrieved February 7, 2020, from//clarivate.libguides.com/woscc/searchtips.

[30] C.A. Scott, T.R. Albrecht, R. De Grenade, A. Zuniga-Teran, R.G. Varady, B. Thapa, Water security and the pursuit of food, energy, and earth systems resilience, Water Int. 43 (8) (2018) 1055–1074, https://doi.org/10.1080/02508060.2018.1534564.

[31] J.I. Uitto, A.M. Duda, Management of transboundary water resources: lessons from international cooperation for conflict prevention, The Geographical Journal 168 (4) (2002) 365–378, https://doi.org/10.1111/j.0016-7398.2002.00062.x.

[32] United Nations. (2013, May 8). What is Water Security? Infographic. Retrieved March 9, 2020, from https://www.unwater.org/publications/water-security-infographic/.

[33] C.J. Vörösmarty, C. Pahl-Wostl, A. Bhaduri, Water in the anthropocene: New perspectives for global sustainability, Curr. Opin. Environ. Sustain. 5 (6) (2013) 535–538, https://doi.org/10.1016/j.cosust.2013.11.011.

[34] A.T. Wolf, Shared waters: Conflict and cooperation, Annu. Rev. Environ. Resour. 32 (2007) 241–269, https://doi.org/10.1146/annurev.energy.32.041006.101434.
[35] A.T. Wolf, A. Kramer, A. Carius, G.D. Dabelko, State of the World: Redefining Global Security, State of the World, 2005, p. 27.
[36] M. Zeitoun, B. Lankford, T. Krueger, T. Forsyth, R. Carter, A.Y. Hoekstra, R. Taylor, O. Varis, F. Cleaver, R. Boelens, L. Swatuk, D. Tickner, C.A. Scott, N. Mirumachi, N. Matthews, Reductionist and integrative research approaches to complex water security policy challenges, Global Environ. Change 39 (2016) 143–154.
[37] E. Babaii, Y. Taase, Author-assigned Keywords in Research Articles: Where Do They Come From, Iran. J. Appl. Linguist. 16 (2) (2013) 1–19.
[38] T. Kuokkanen, Water Security and International Law, Potchefstroom Electron. Law J. 20 (2017) 1–22, https://doi.org/10.17159/1727-3781/2017/v20i0a1652.
[39] M. Heckner, S. Mühlbacher, C. Wolff, Tagging tagging. Analysing user keywords in scientific bibliography management systems. J. Digit. Inf. (2008).