Review

Understanding Recent Trends in Global Sustainable Development Goal 6 Research: Scientometric, Text Mining and an Improved Framework for Future Research

Ajishnu Roy [1]**, Aman Basu** [2]**, Yanyu Su** [1]**, Yan Li** [1,3,]*** and Xuhui Dong** [1,3,]*****

<u>**Supplementary Information**</u>

<u>Search Criteria</u>: Date 25th March 2021. From WoS core collections

TS: "SDG 6" OR "SDG6" OR "sustainable development Goal 6"

Timespan: All years.

Indexes: SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC.

Search results: 299

Removed: 10 [6 (weak connection to SDG 6) + 1 (duplicate) + 2 (not in English language) + 1 (unavailable full text)]

Note on keywords:

Since it is easily understandable that some commonly used terms are generally used most predominantly, but without yielding any specific characteristic insights, we have tried to omit or avoid these few related terms from keyword analysis as much as possible. These terms would be – 'water', 'sanitation', sustainable development goals', 'sustainable development goal', 'sustainable development goal 6', 'sdgs', 'sdg', 'sdg 6', 'sdg6', 'development goals', 'goals', 'sustainability', 'sustainable', 'study', 'indicators', 'indicator', 'analysis' etc.
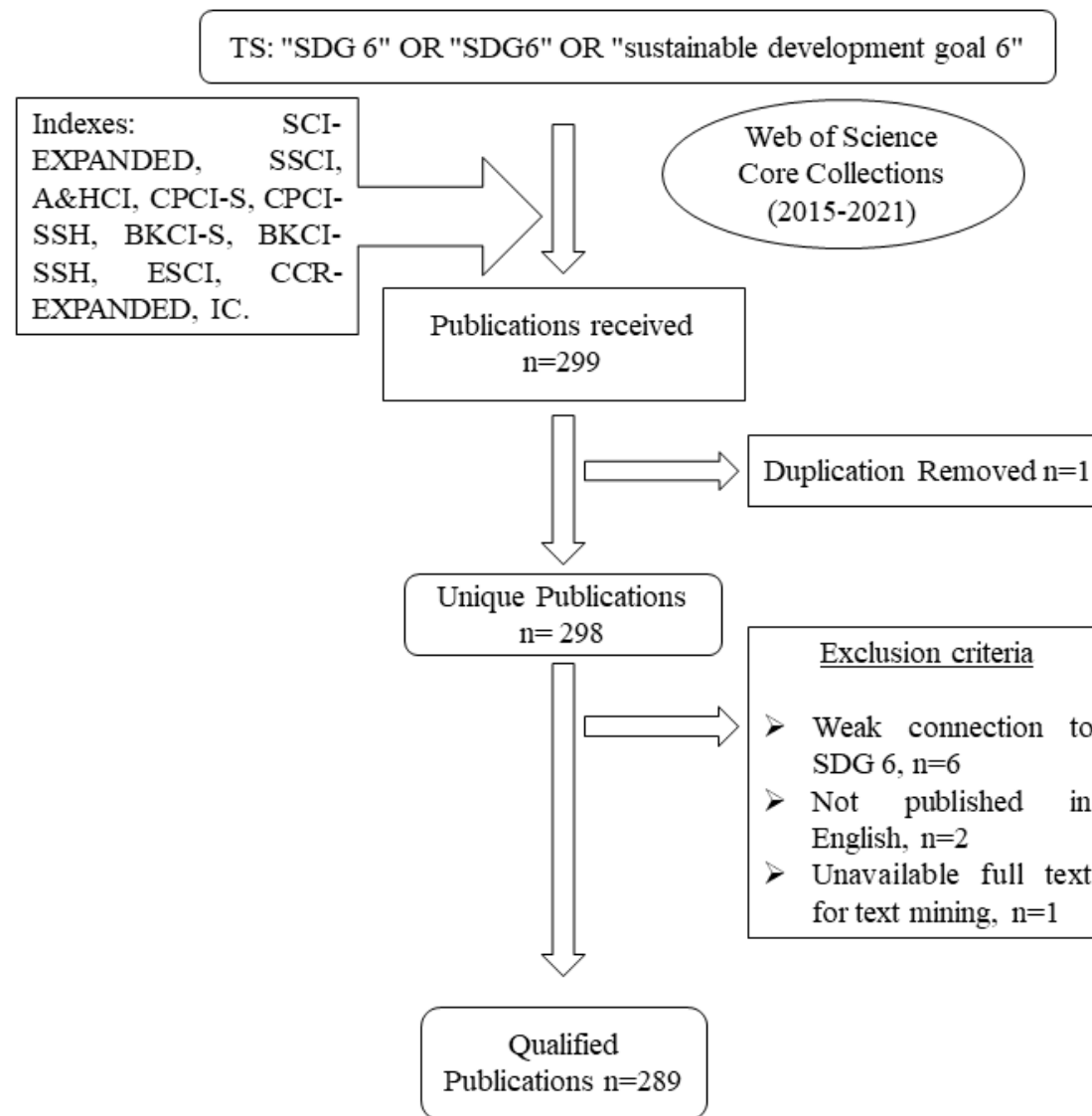
Figure. Flow chart of bibliometric search strategy.

Table S1. Targets of UN SDG 6.

Target	Focus	Deadline
6.1	achieve universal and equitable access to safe and affordable drinking water for all	2030
6.2	achieve access to adequate and equitable sanitation and hygiene for all and end open defecation, paying special attention to the needs of women and girls and those in vulnerable situations	
6.3	improve water quality by reducing pollution, eliminating dumping and minimizing release of hazardous chemicals and materials, halving the proportion of untreated wastewater and substantially increasing recycling and safe reuse globally	
6.4	substantially increase water-use efficiency across all sectors and ensure sustainable withdrawals and supply of freshwater to address water scarcity and substantially reduce the number of people suffering from water scarcity	
6.5	implement integrated water resources management at all levels, including through transboundary cooperation as appropriate	
6.6	protect and restore water-related ecosystems, including mountains, forests, wetlands, rivers, aquifers and lakes	2020 (i.e., deadline already crossed)
6.a	expand international cooperation and capacity-building support to developing countries in water- and sanitation-related activities and programmes, including water harvesting, desalination, water efficiency, wastewater treatment, recycling and reuse technologies	2030
6.b	Support and strengthen the participation of local communities in improving water and sanitation management	NA

Topic model

To obtain the optimal number of topics suitable for the dataset, we tested it using "Griffiths2004"[1], "CaoJuan2009"[2], "Arun2010"[3], "Deveaud2014"[4] methods for a full range of 2 to 80 topics. After obtaining the desired number of topics, the latent Dirichlet allocation (LDA) method was used for constructing the topic model.

References

1. Rajkumar Arun, V. Suresh, C. E. Veni Madhavan, and M. N. Narasimha Murthy. 2010. On finding the natural number of topics with latent dirichlet allocation: Some observations. In Advances in knowledge discovery and data mining, Mohammed J. Zaki, Jeffrey Xu Yu, Balaraman Ravindran and Vikram Pudi (eds.). Springer Berlin Heidelberg, 391–402. http://doi.org/10.1007/978-3-642-13657-3_43

2. Cao Juan, Xia Tian, Li Jintao, Zhang Yongdong, and Tang Sheng. 2009. A density-based method for adaptive lda model selection. Neurocomputing — 16th European Symposium on Artificial Neural Networks 2008 72, 7–9: 1775–1781. http://doi.org/10.1016/j.neucom.2008.06.011

3. Romain Deveaud, Éric SanJuan, and Patrice Bellot. 2014. Accurate and effective latent concept modeling for ad hoc information retrieval. Document numérique 17, 1: 61–84. http://doi.org/10.3166/dn.17.1.61-84

4. Thomas L. Griffiths and Mark Steyvers. 2004. Finding scientific topics. Proceedings of the National Academy of Sciences 101, suppl 1: 5228–5235. http://doi.org/10.1073/pnas.0307752101

Terms omitted for text mining

For Bigram plot: "et", "al", "can", "also", "may", "s", "m", "will", "c", "e", "however", "b", "p", "j", "n", "t", "r", "fig", "d", "66", "a", "l", "n", "f", "ie", "g", "table", "the", "in", "this", "used", "using", "per", "within", "doi", "x", "httpsdoiorg", "peer", "review", "online". All numbers (0-9) deleted.

For Topic model: "et", "al", "can", "also", "may", "s", "m", "will", "c", "e", "however", "b", "p", "j", "n", "t", "r", "fig", "d", "66", "a", "l", "n", "f", "ie", "g", "table", "the", "in", "this", "used", "using", "per", "within", "doi", "x", "httpsdoiorg", "peer", "review", "online", "water", "6", "sdg", "sdgs", "figure", "use", "study", "data", "bosnia", "herzegovina", "sustainable", "development", "goal", "goals". Numbers deleted: 0-9 (except 6).
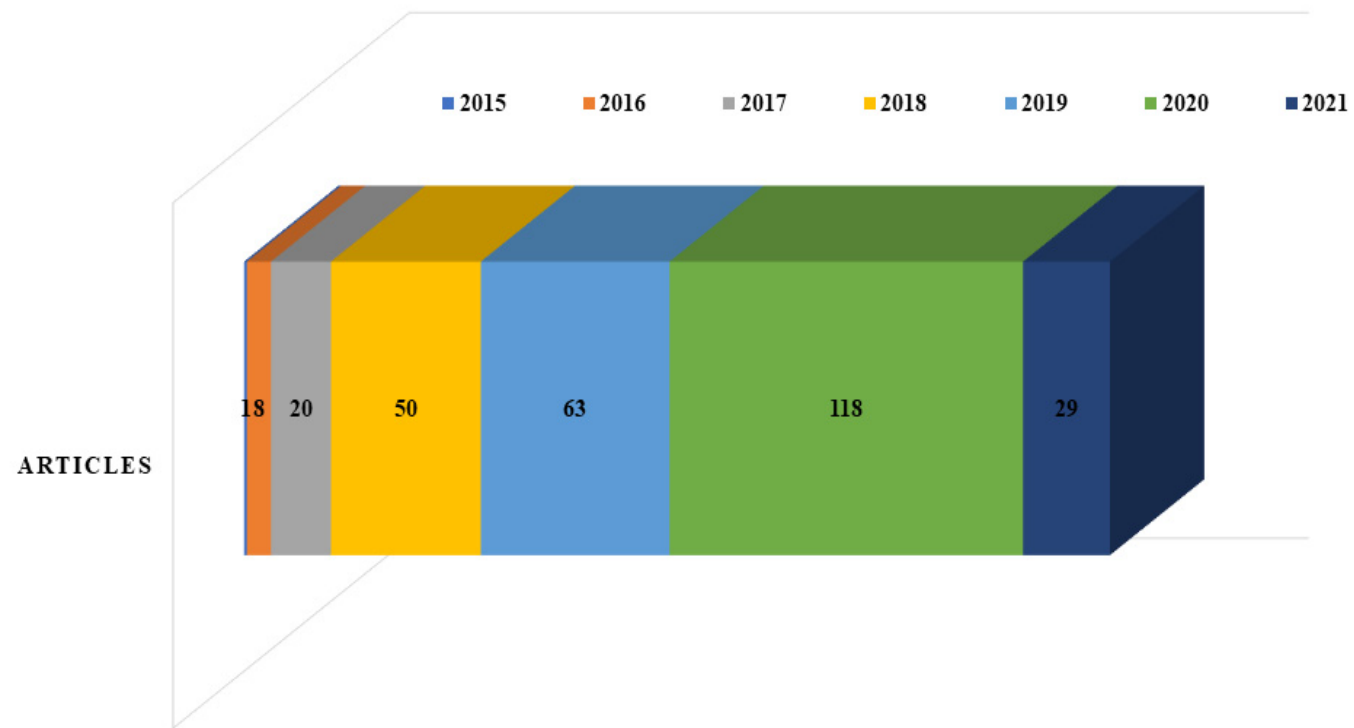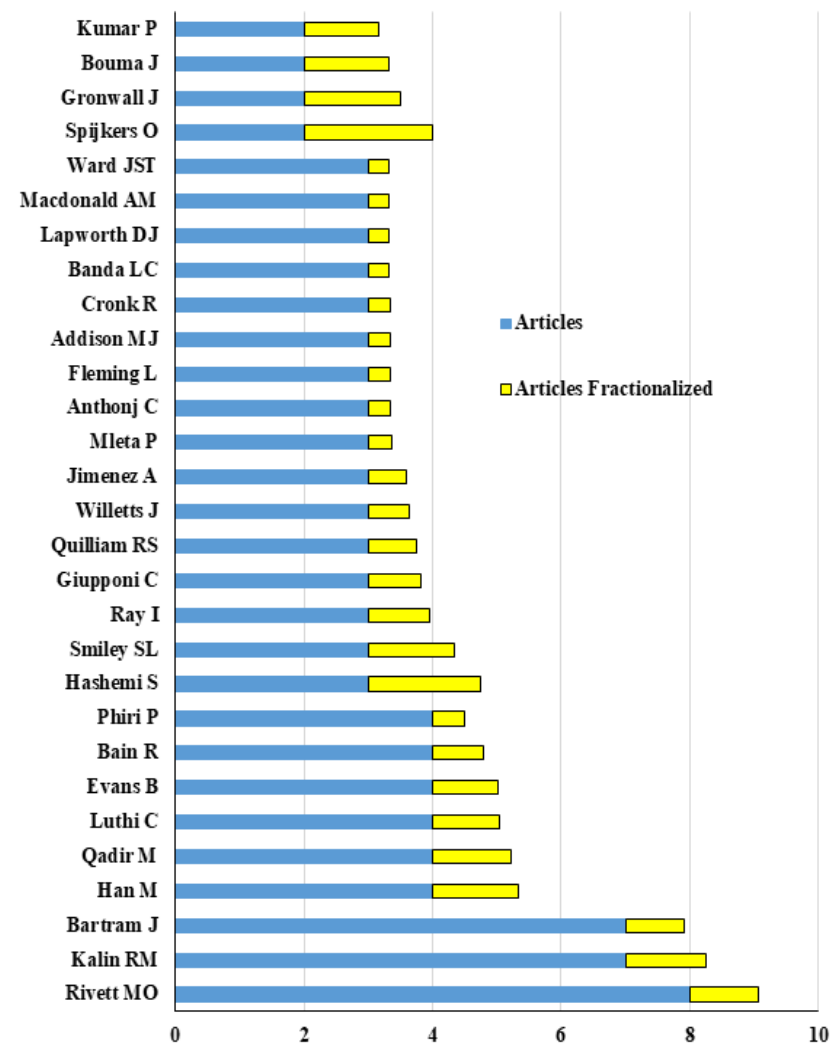
Figure S1. Annual scientific production trends.
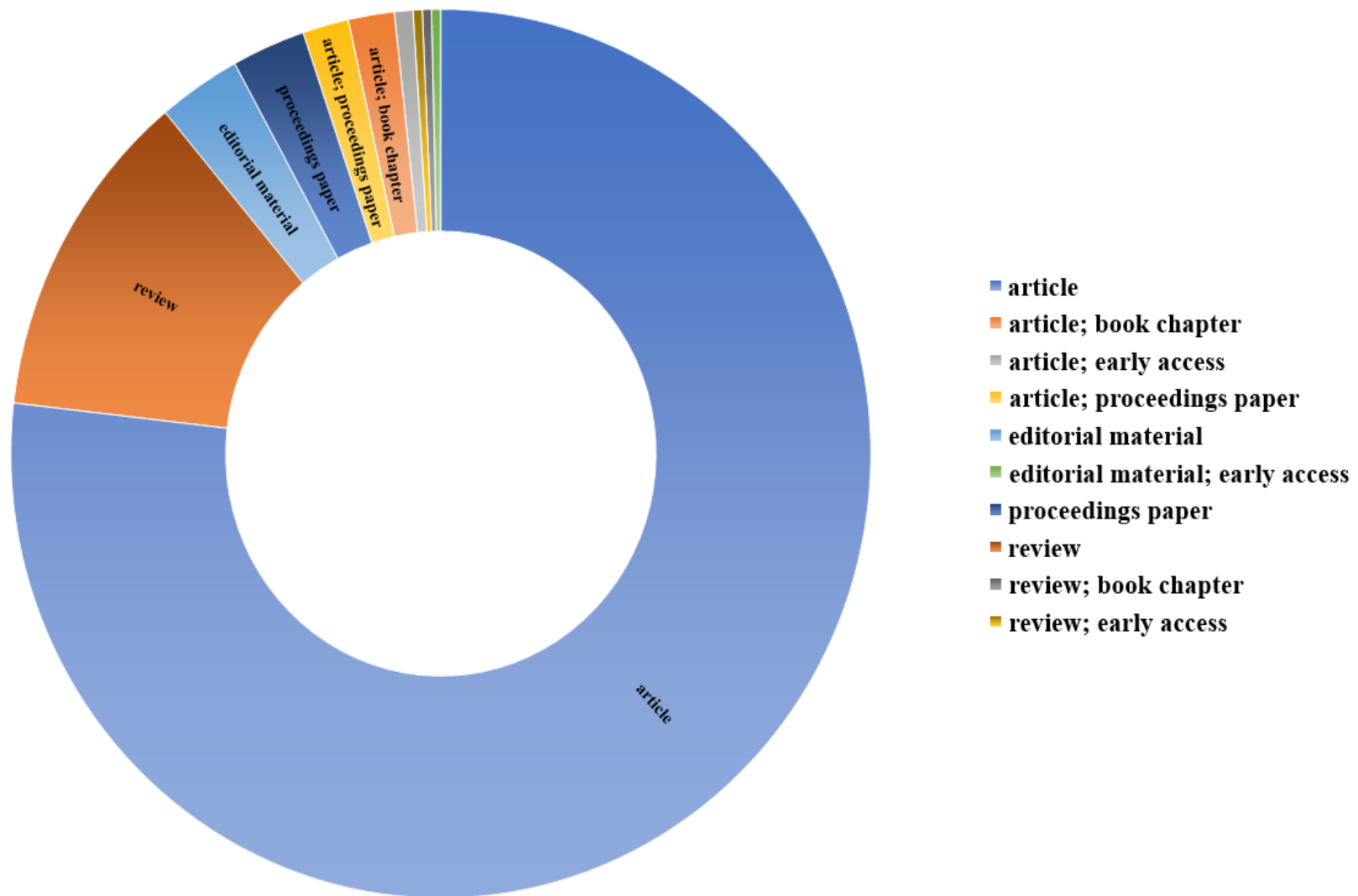
Figure S2. Annual scientific production trends.
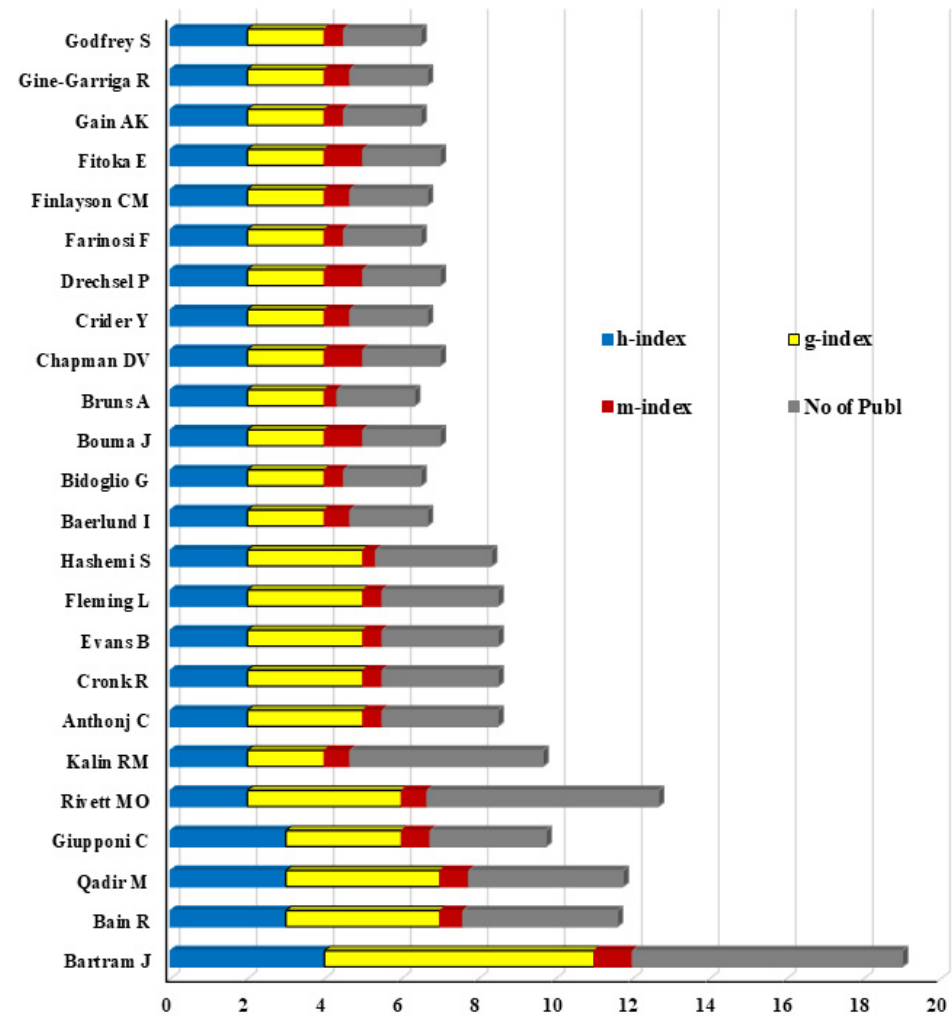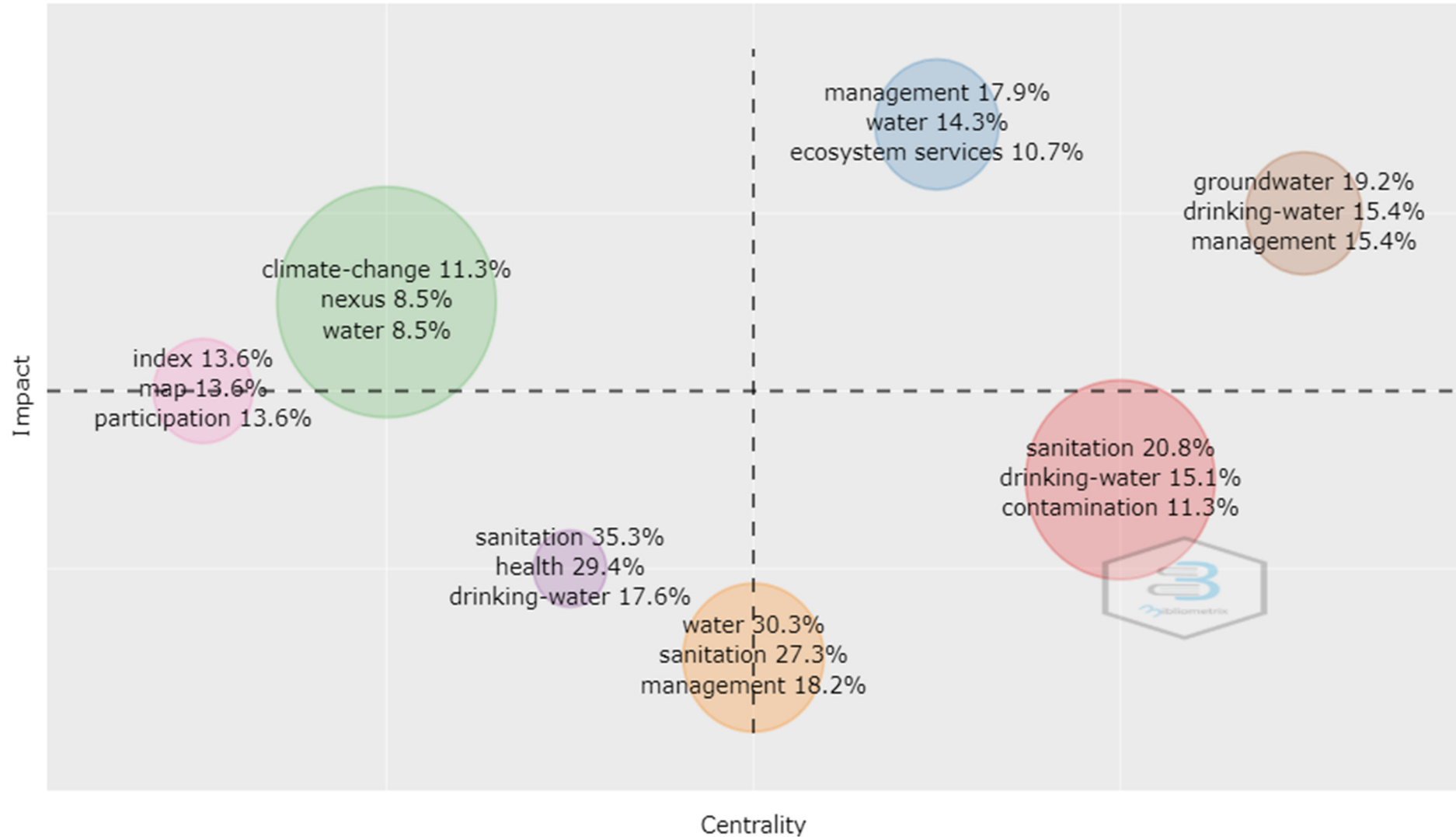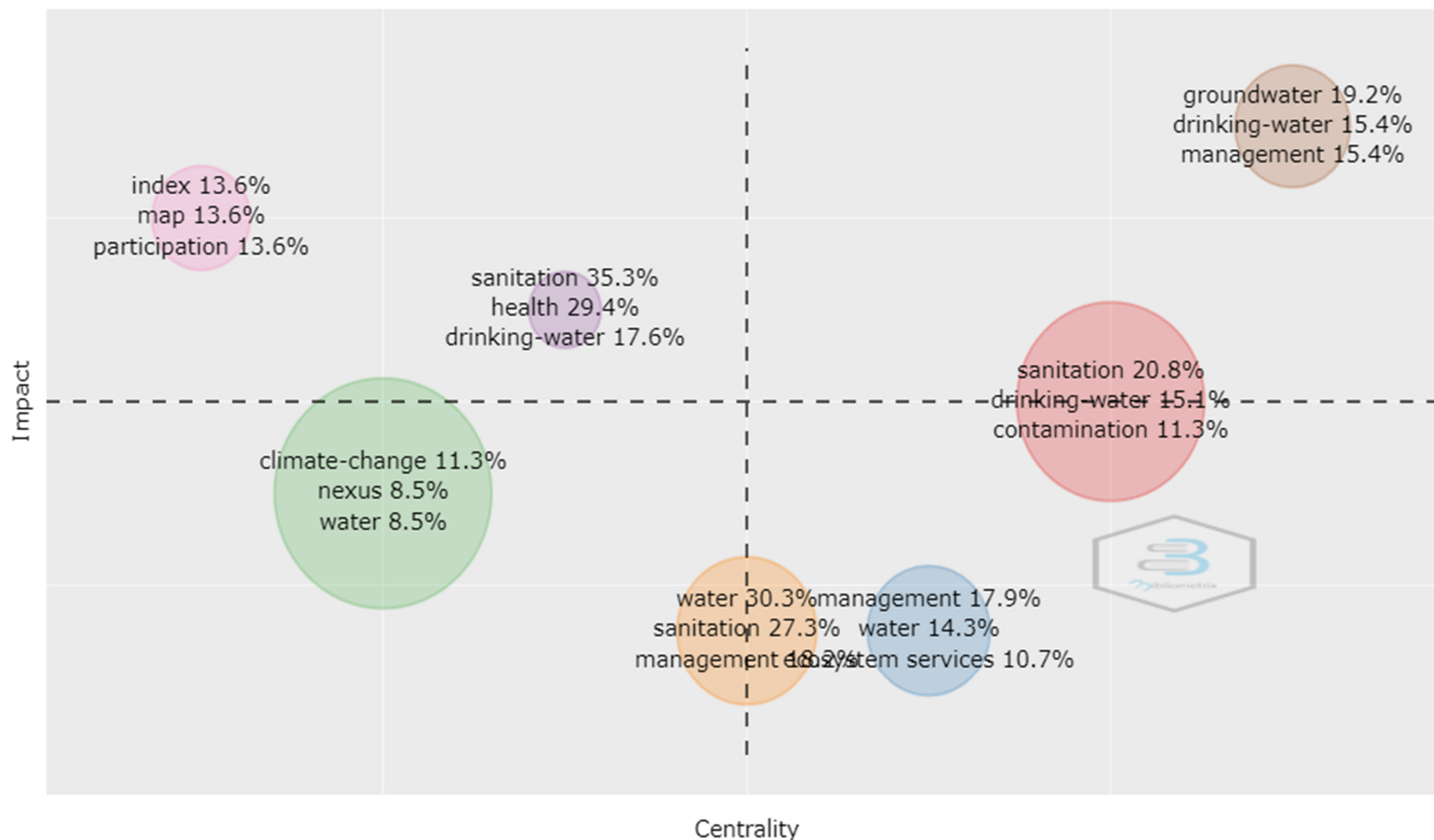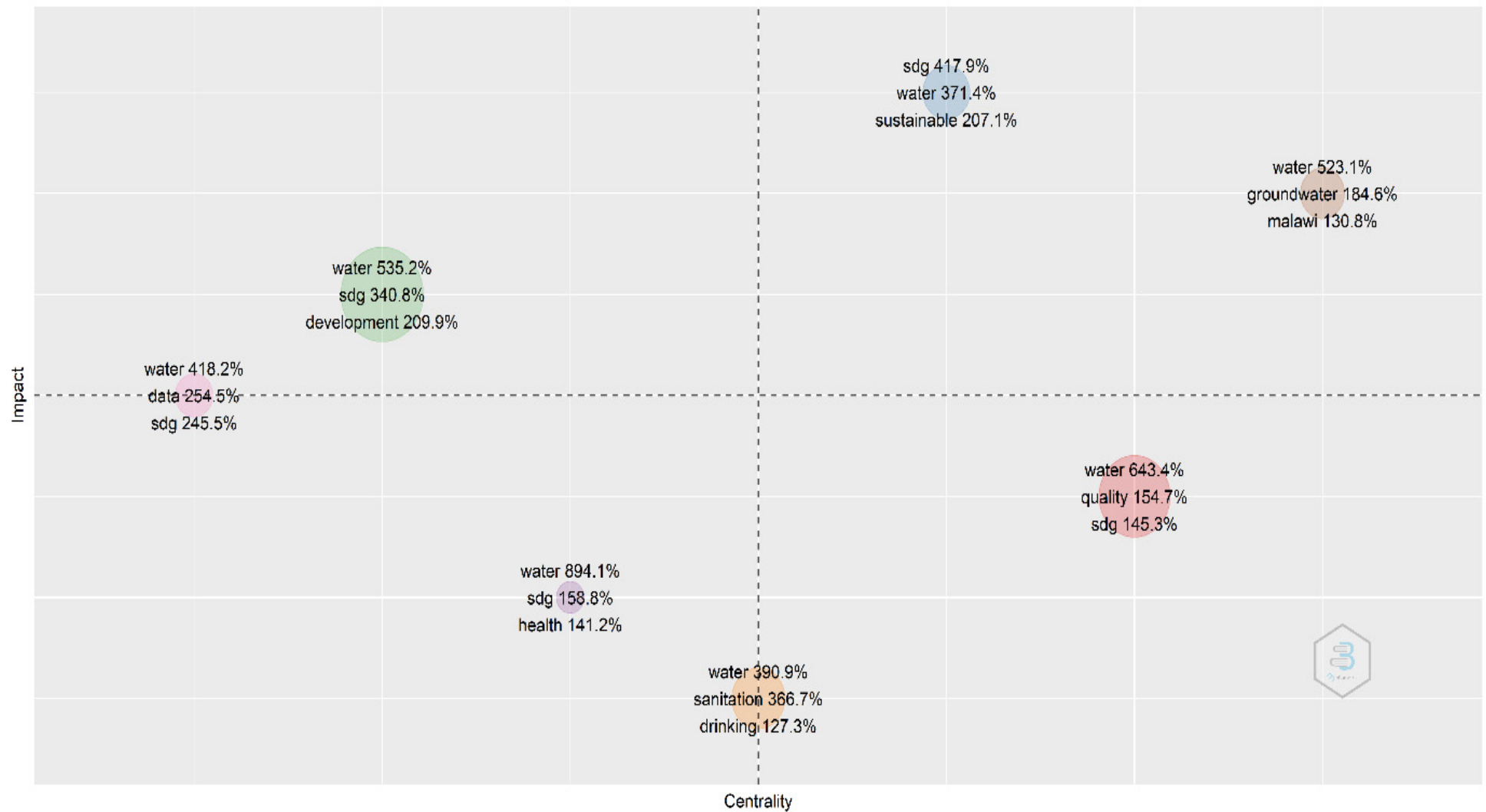
Figure S3. Most relevant authors.

Figure S4. Authors' Local Impact.

Figure S5. Clusters by documents coupling with global citation (GC) scores of global SDG 6 research (as per Keywords Plus).

Figure S6. Clusters by documents coupling with local citation (LC) scores of global SDG 6 research (as per Keywords Plus).
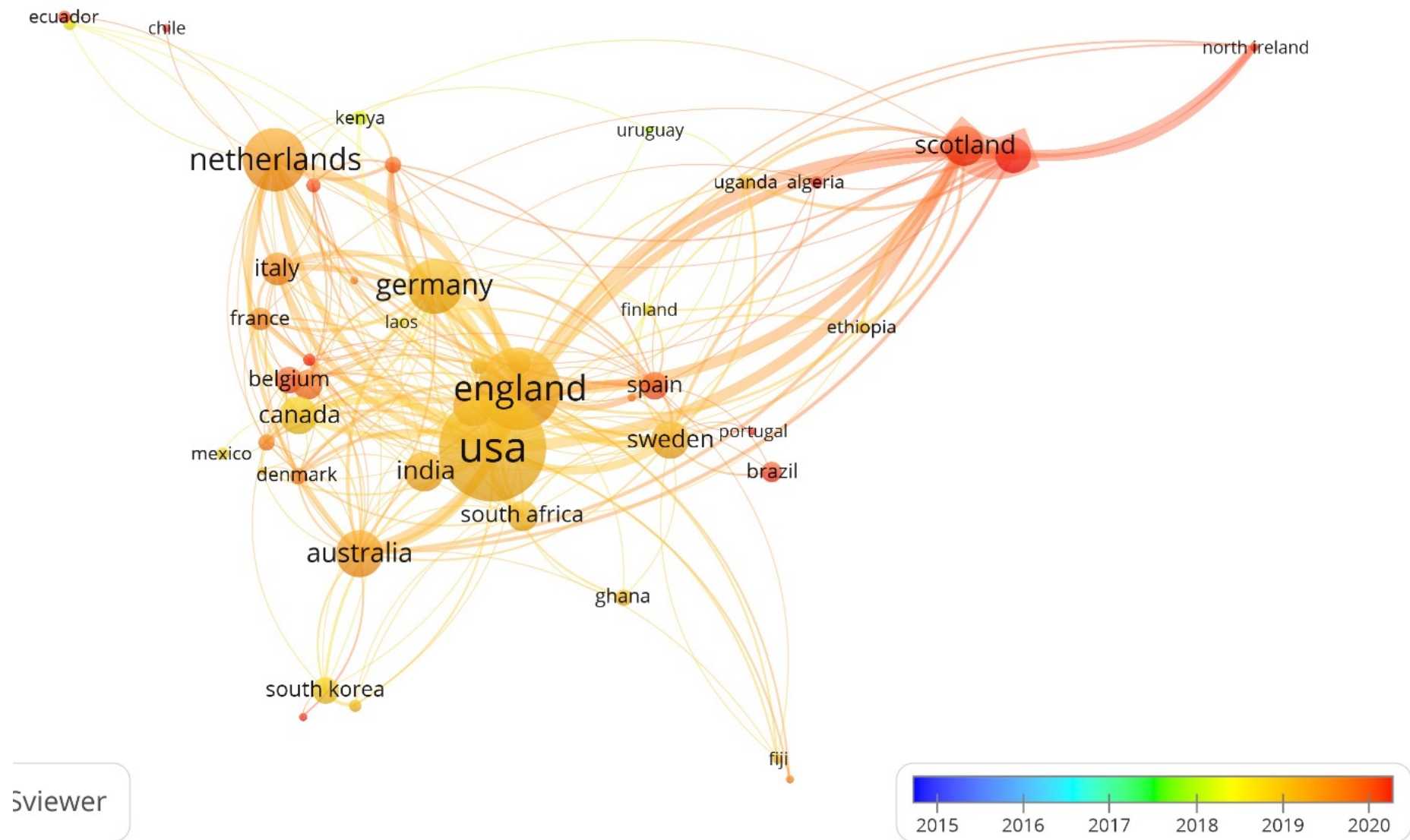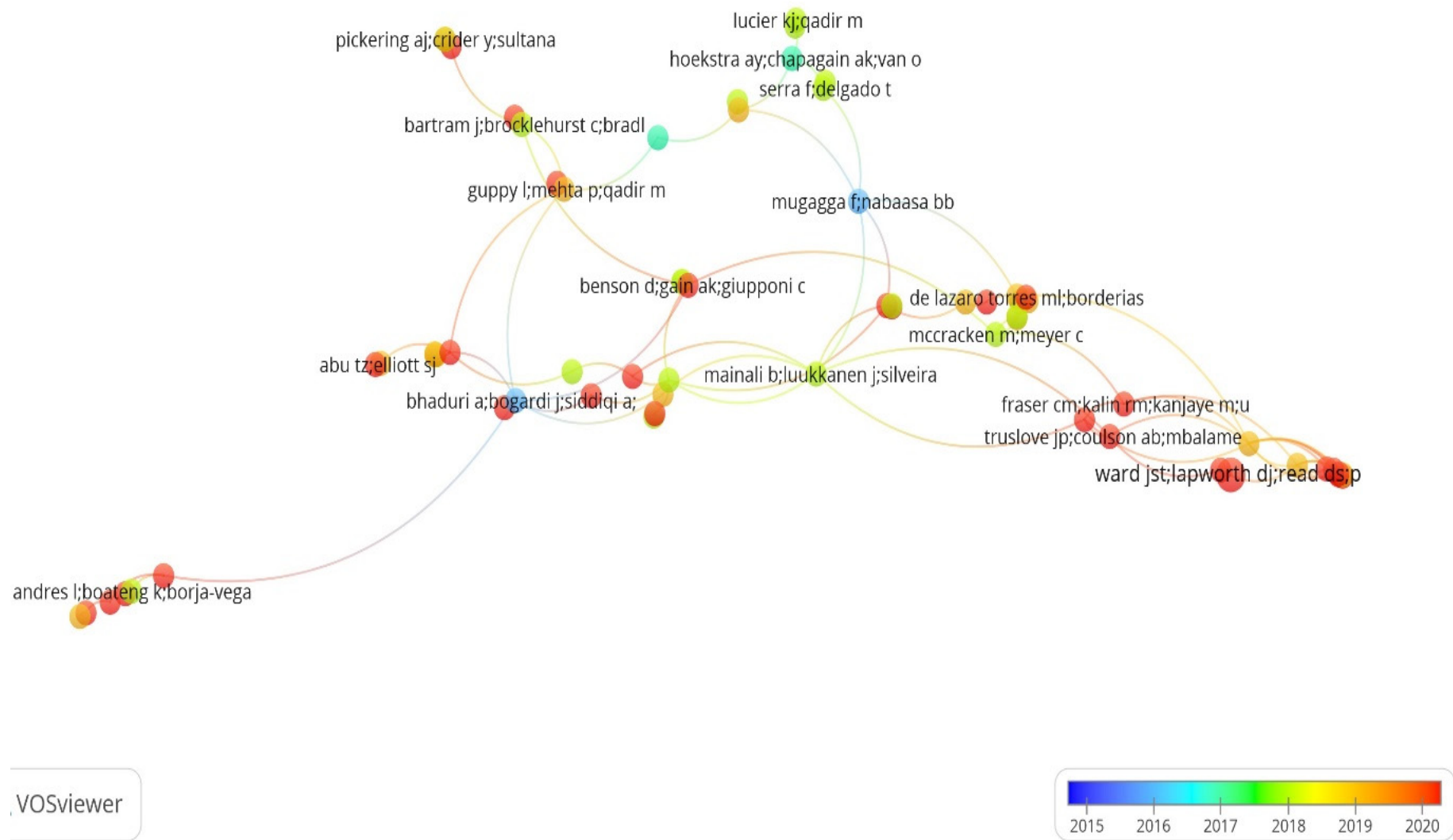
Figure S7. Clusters by documents coupling with global citation (GC) scores of global SDG 6 research (as per content of abstracts).
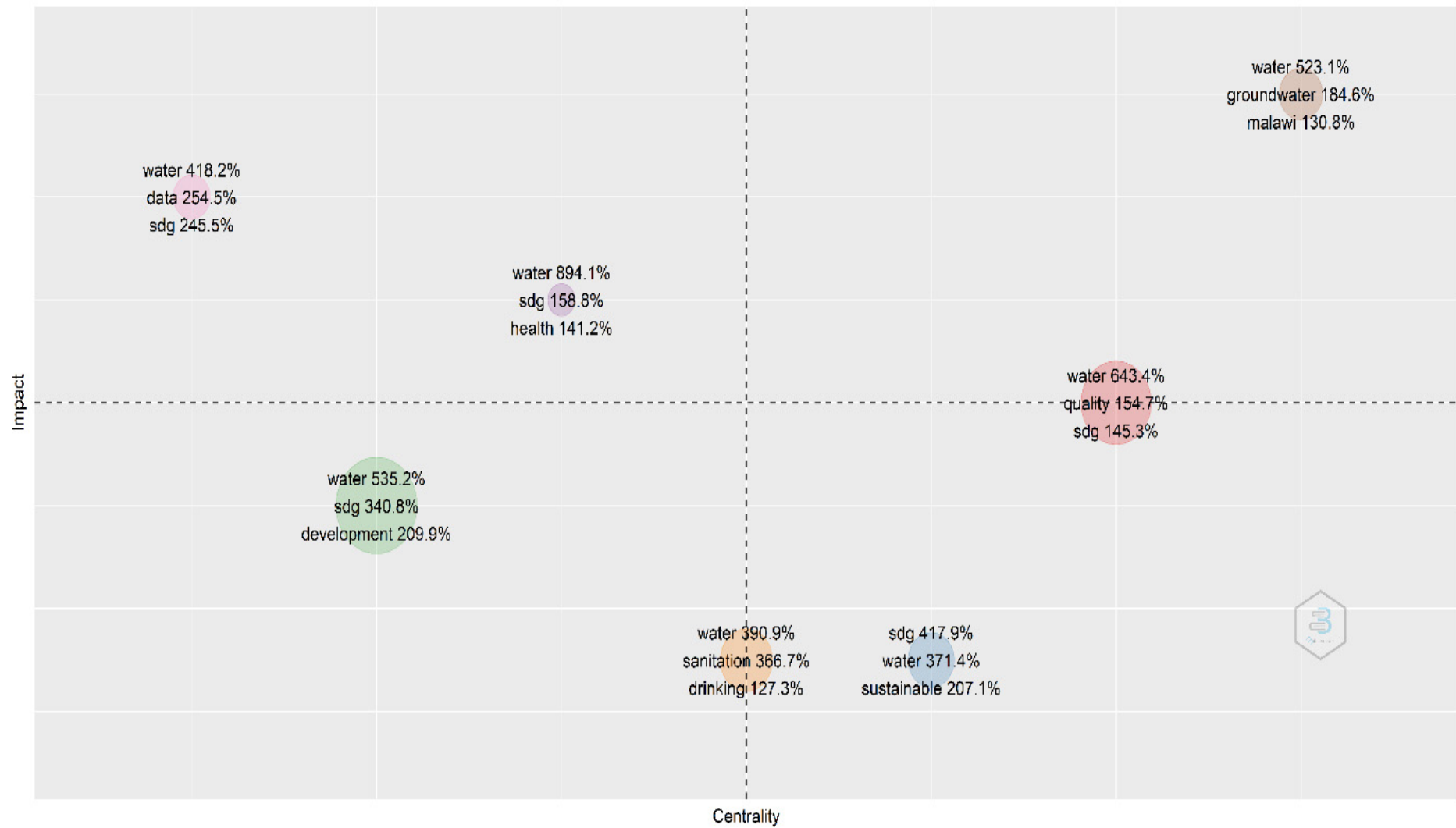
Figure S8. Clusters by documents coupling with local citation (LC) scores of global SDG 6 research (as per content of abstracts).

Figure S9. Citation network of countries of global SDG 6 research.

Figure S10. Citation network of authors of global SDG 6 research.
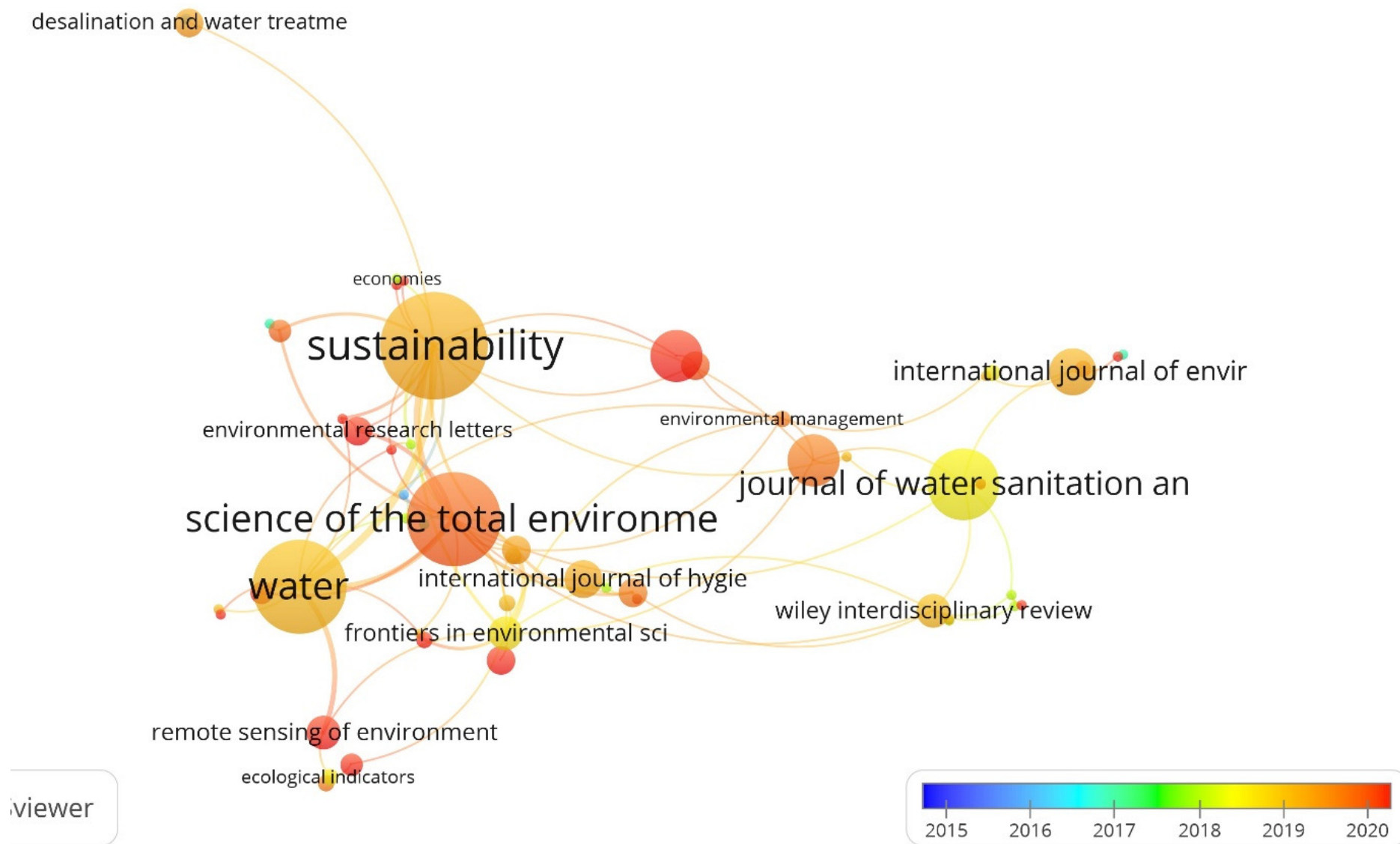
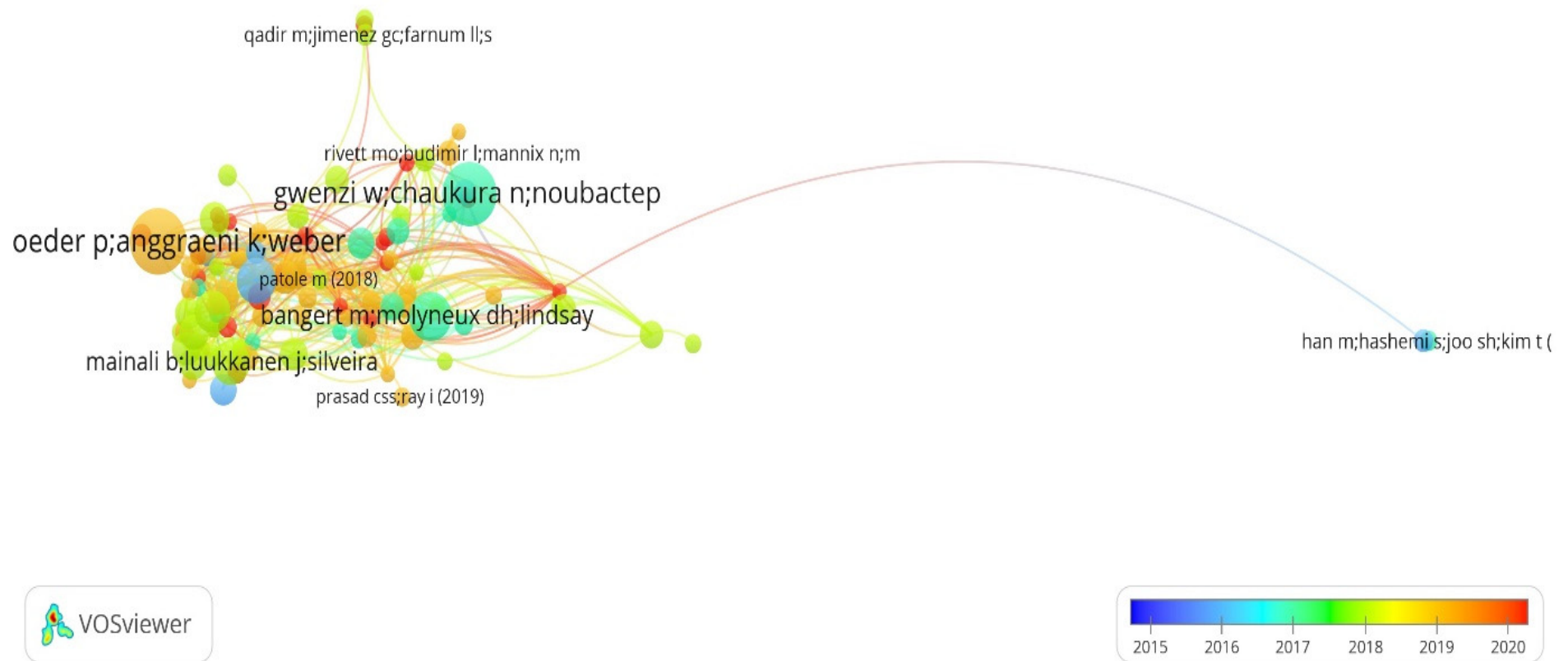**Figure S11.** Citation network of sources of global SDG 6 research.
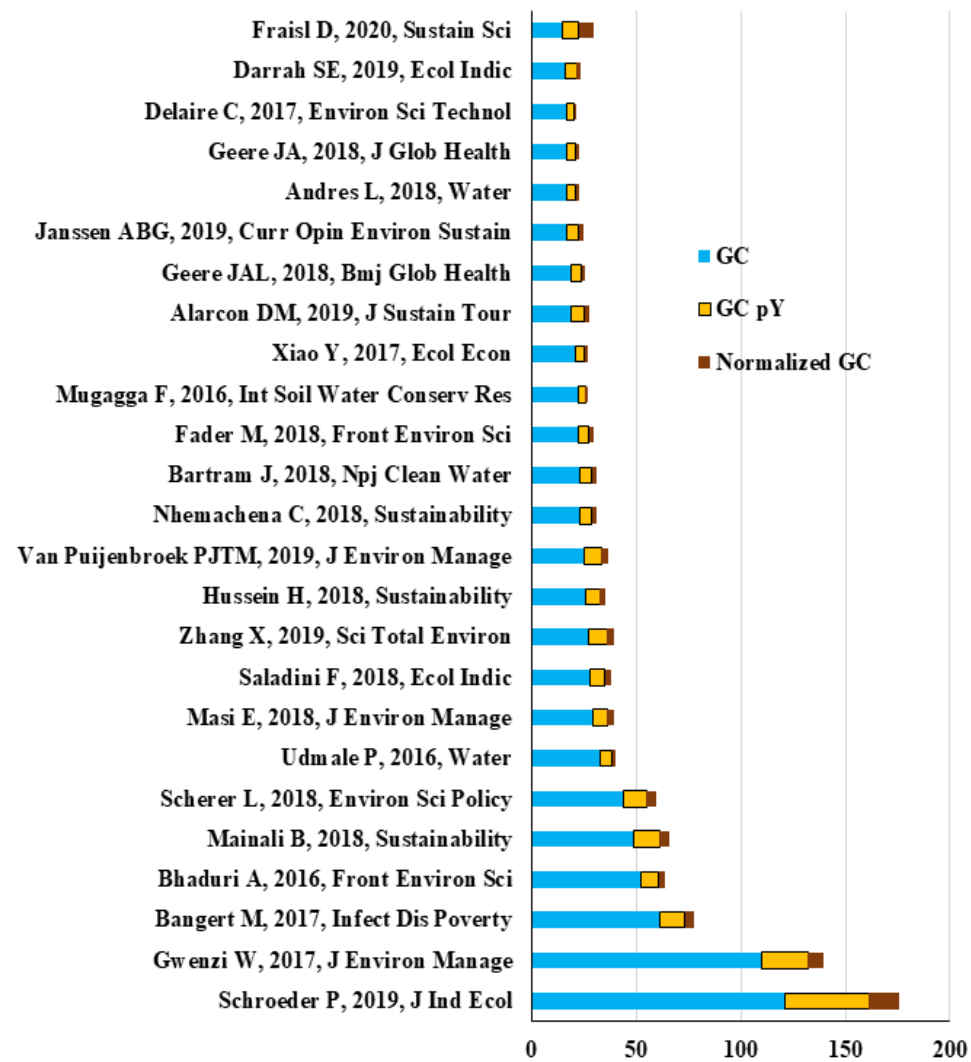
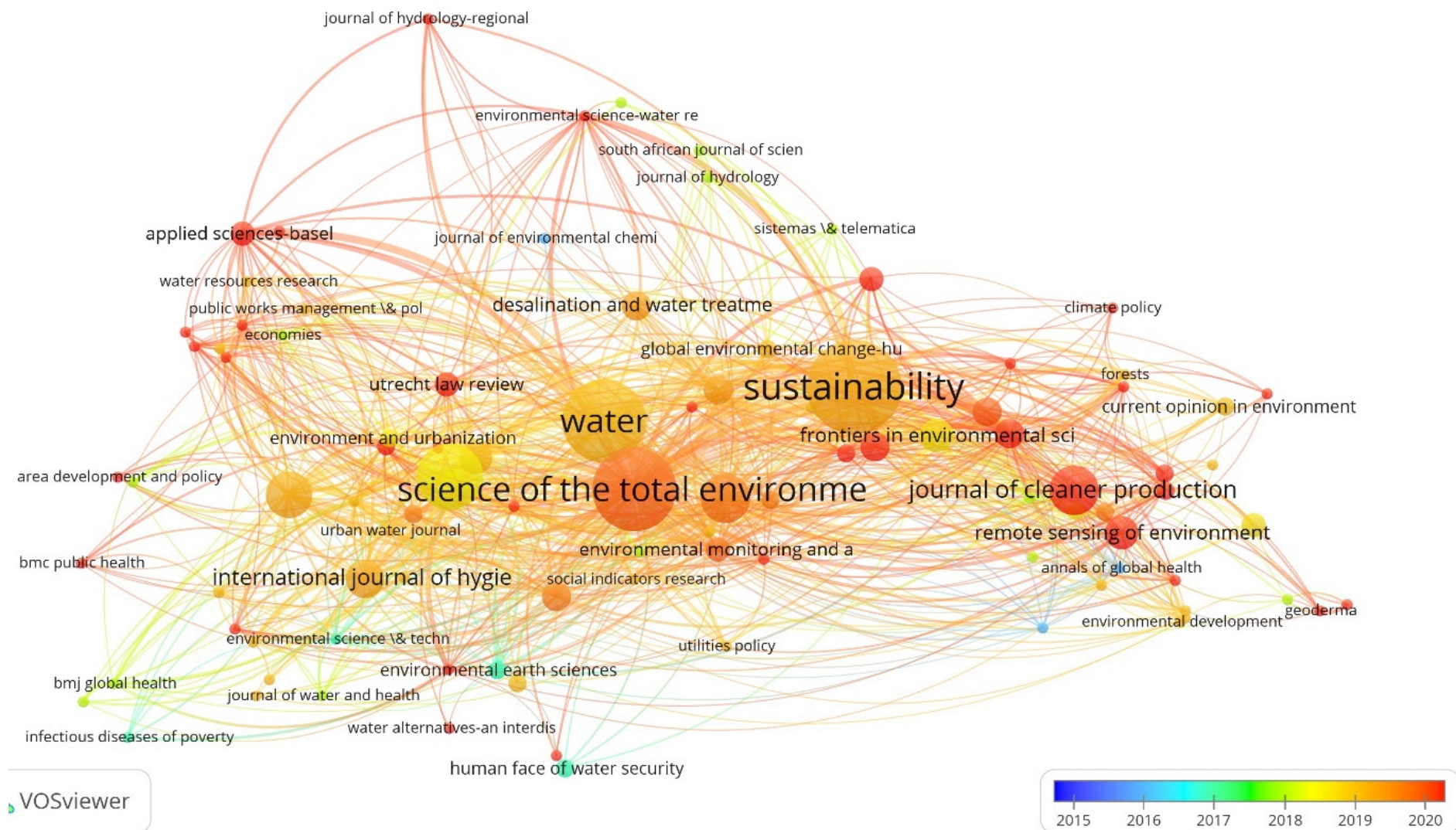**Figure S12.** Most global cited (GC) documents.

Figure S13. Bibliographic coupling network of authors of global SDG 6 research.
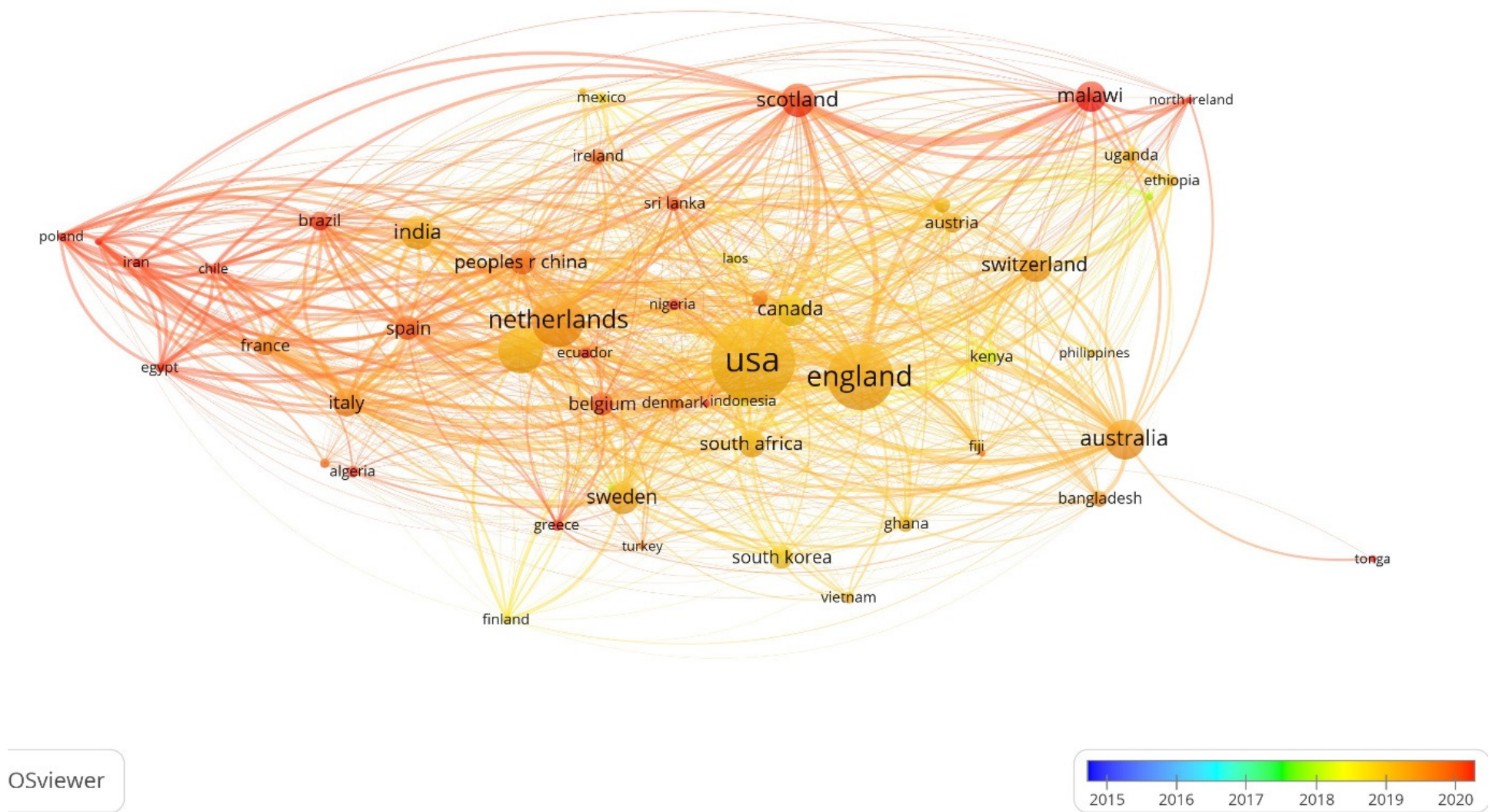
Figure S14. Bibliographic coupling network of countries of global SDG 6 research.

qadir m;jimenez gc;farnum ll;s

rivett mo;budimir l;mannix n;m

gwenzi w;chaukura n;noubactep

oeder p;anggraeni k;weber

patole m (2018)

bangert m;molyneux dh;lindsay

mainali b;luukkanen j;silveira

prasad css;ray i (2019)

han m;hashemi s;joo sh;kim t (

2015 2016 2017 2018 2019 2020

VOSviewer

Figure S15. Bibliographic coupling network of documents of global SDG 6 research.

Figure S16. Bibliographic coupling network of sources of global SDG 6 research.

Figure S17. Word cloud of most frequent keywords in global SDG 6 research (as per Keywords Plus).

Figure S18. Word cloud of most frequent keywords in global SDG 6 research (as per titles).

Figure S19. Word cloud of most frequent keywords in global SDG 6 research (as per content of abstracts).
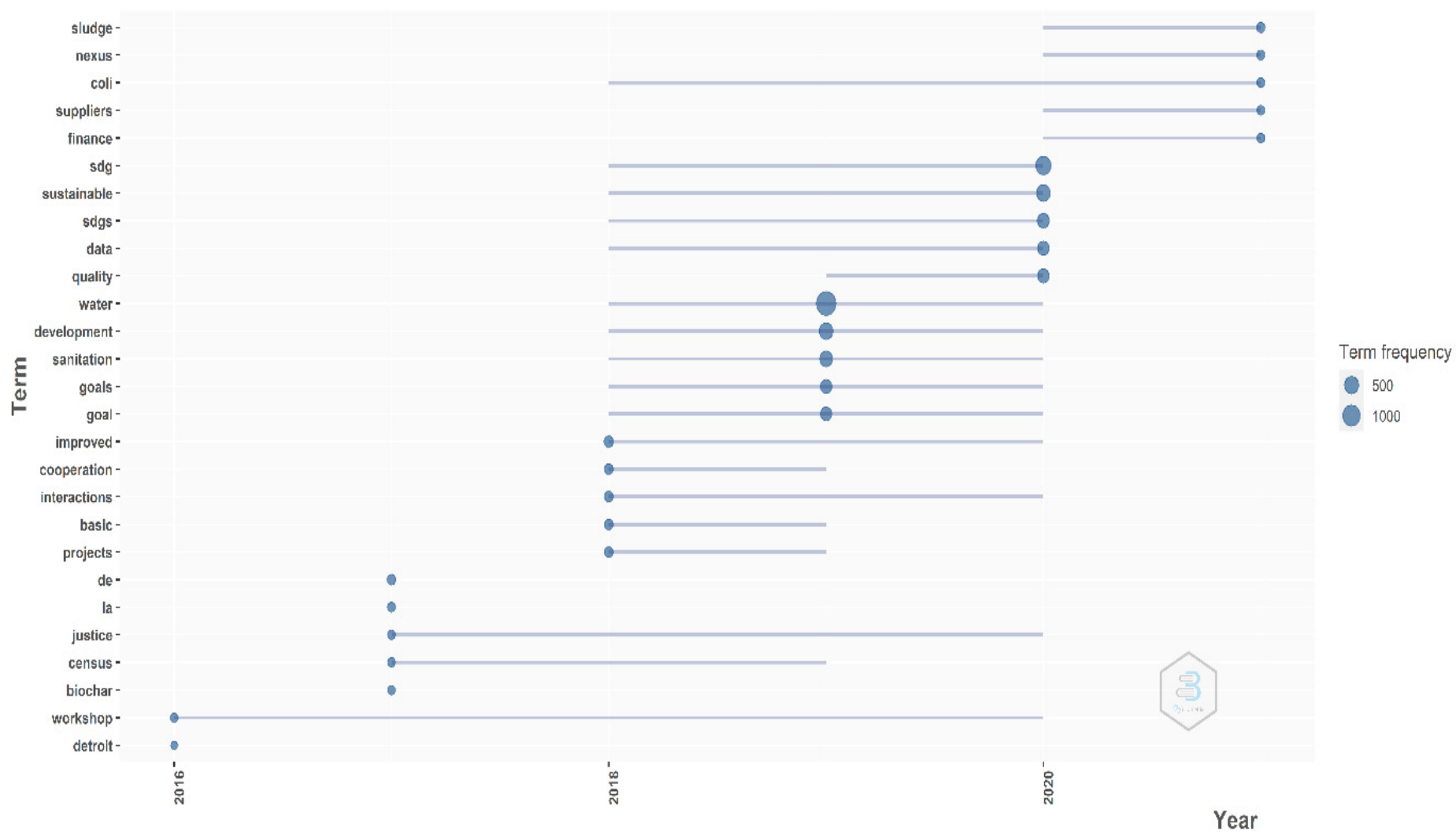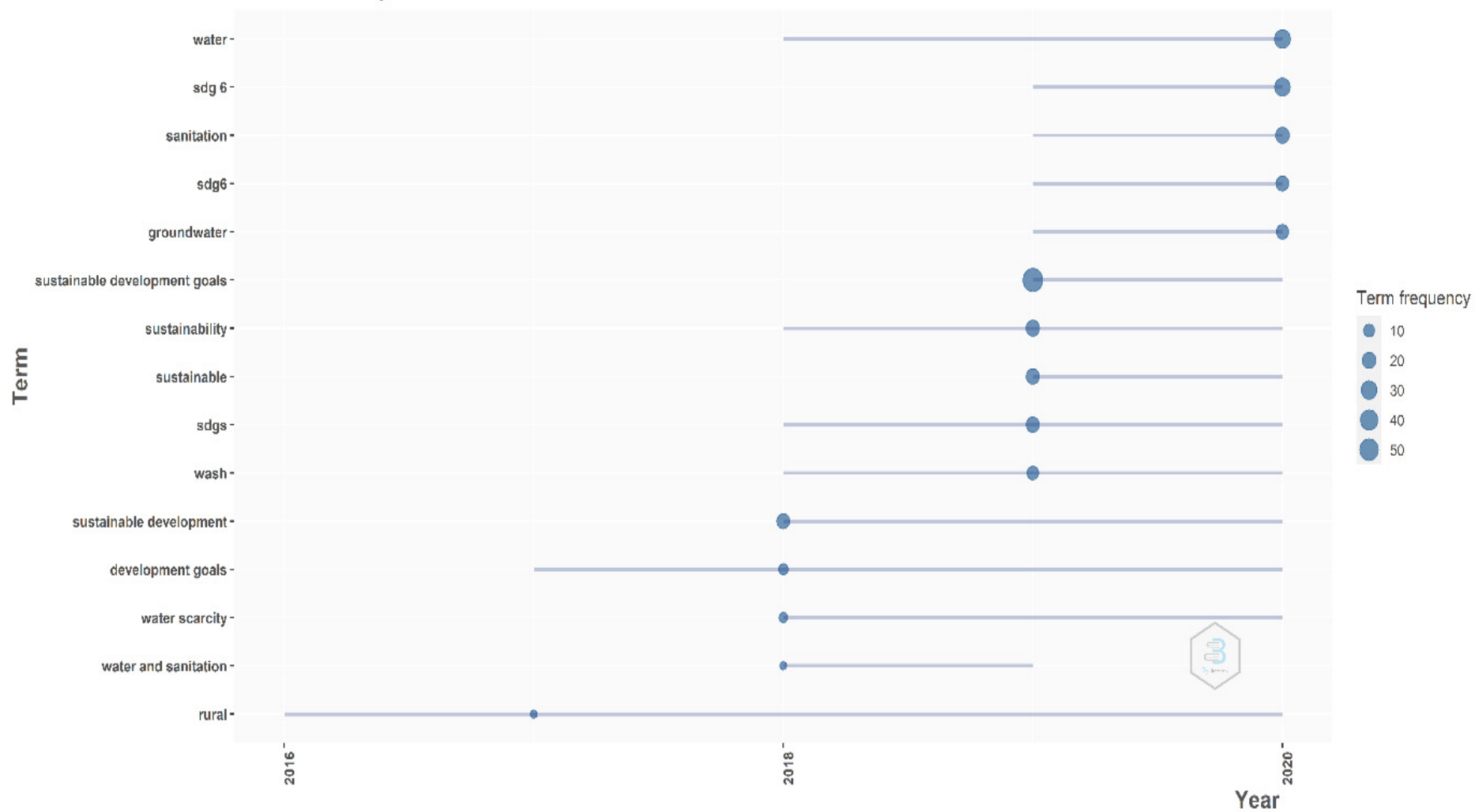
Figure S20. Word cloud of most frequent keywords in global SDG 6 research (as per authors' keywords).
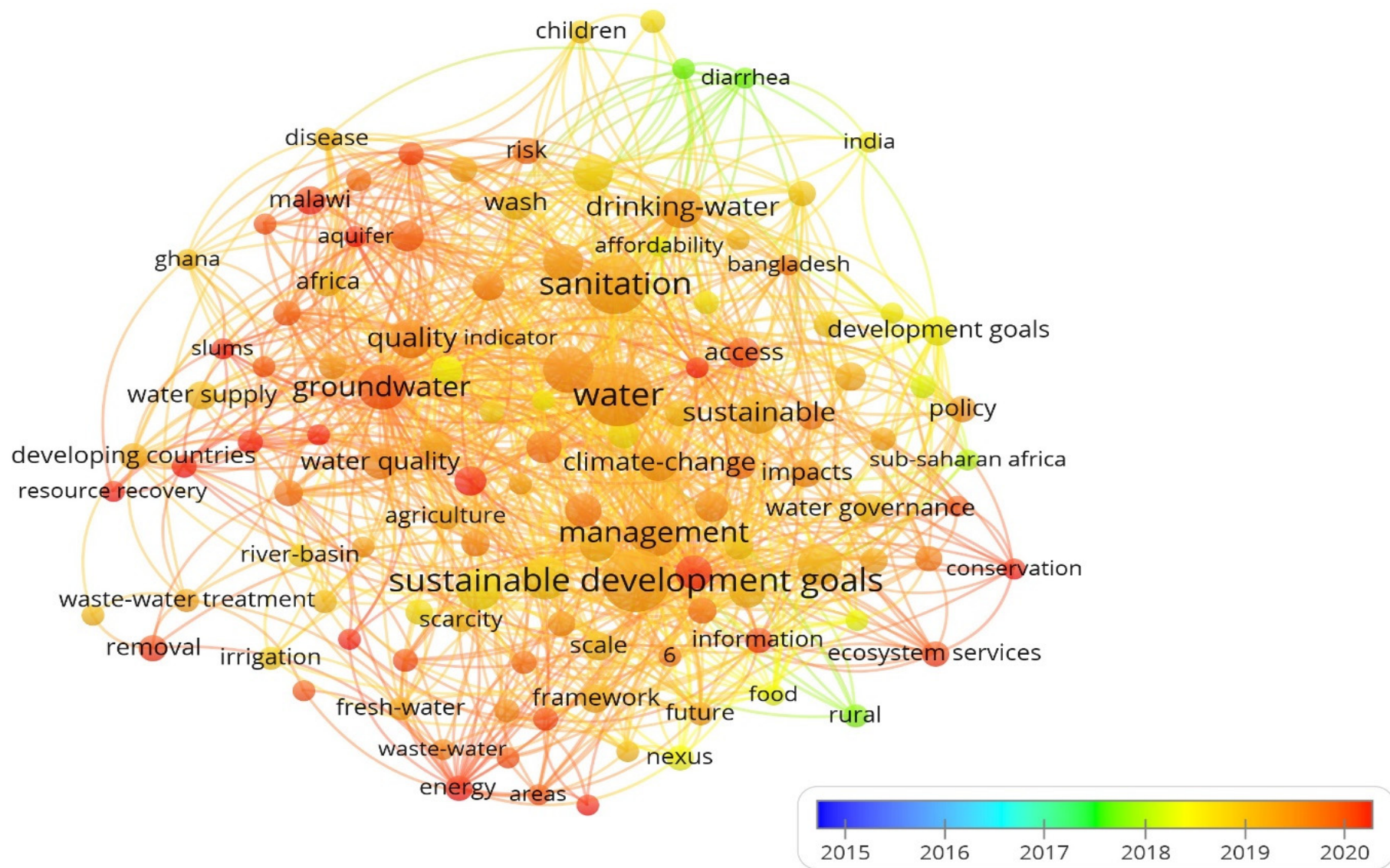
Figure S21. Co-occurrence network of all keywords of global SDG 6 research.

Figure S22. Trending topics of global SDG 6 research (as per content of abstracts).
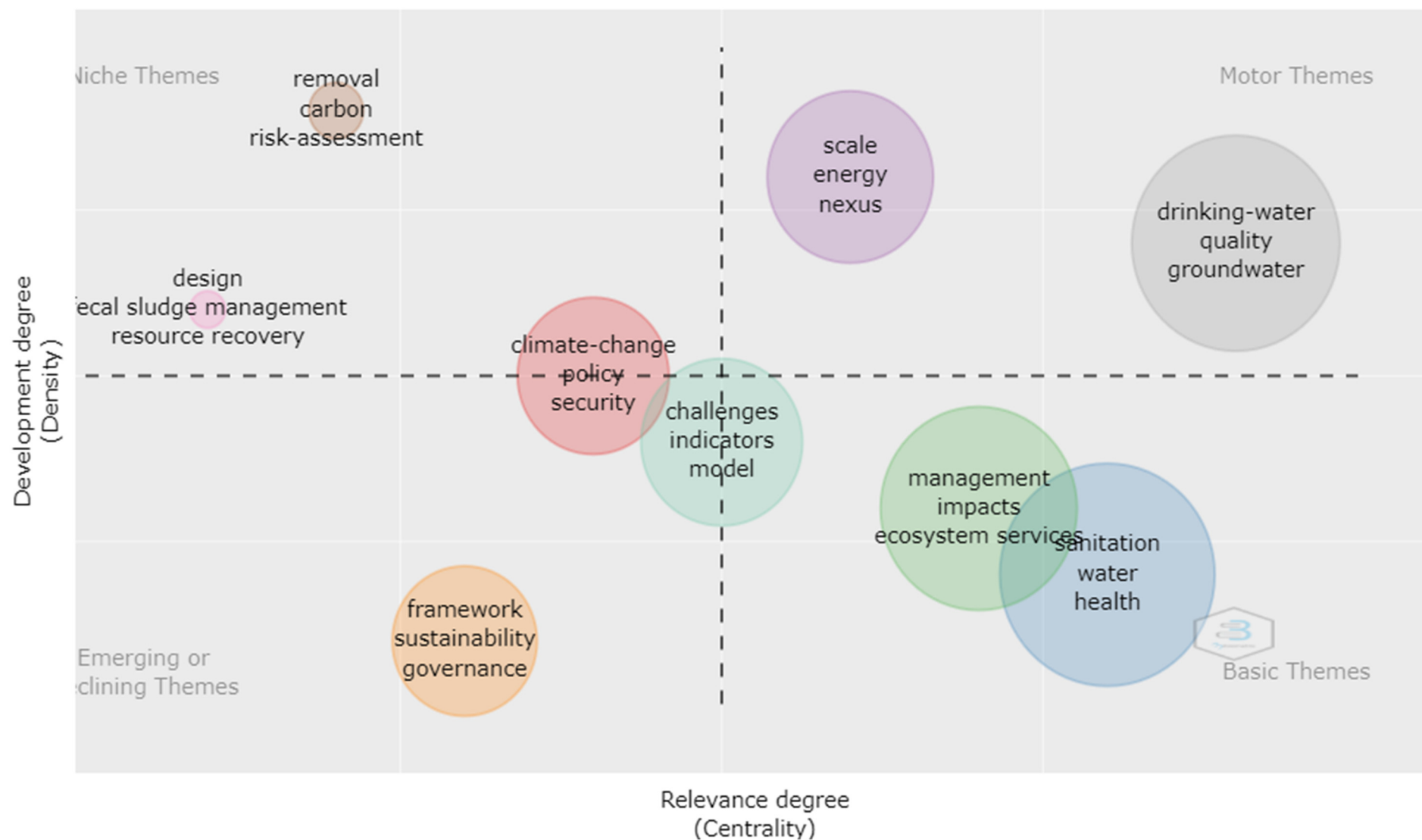
Figure S23. Trending topics of global SDG 6 research (as per authors' keywords).
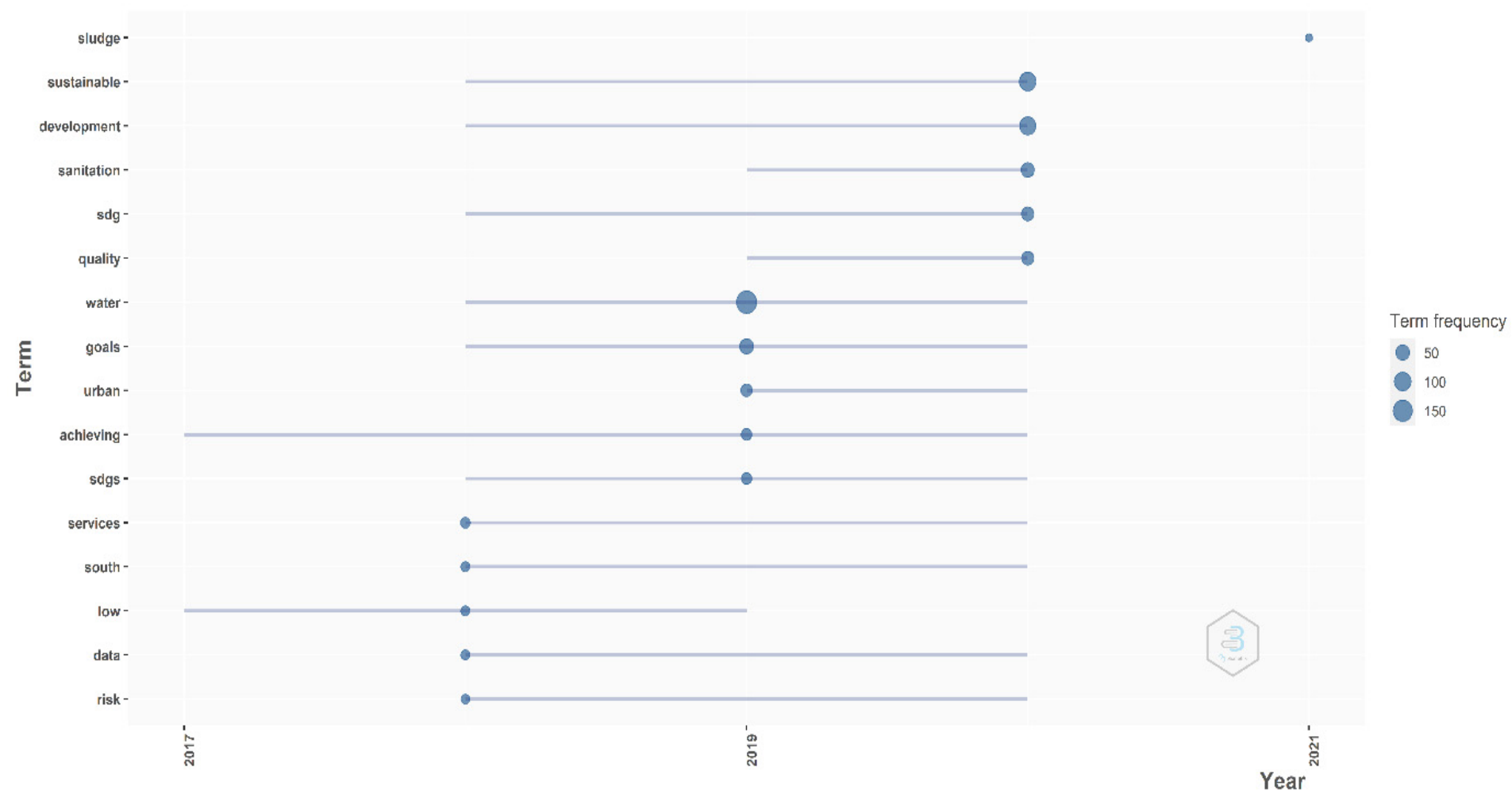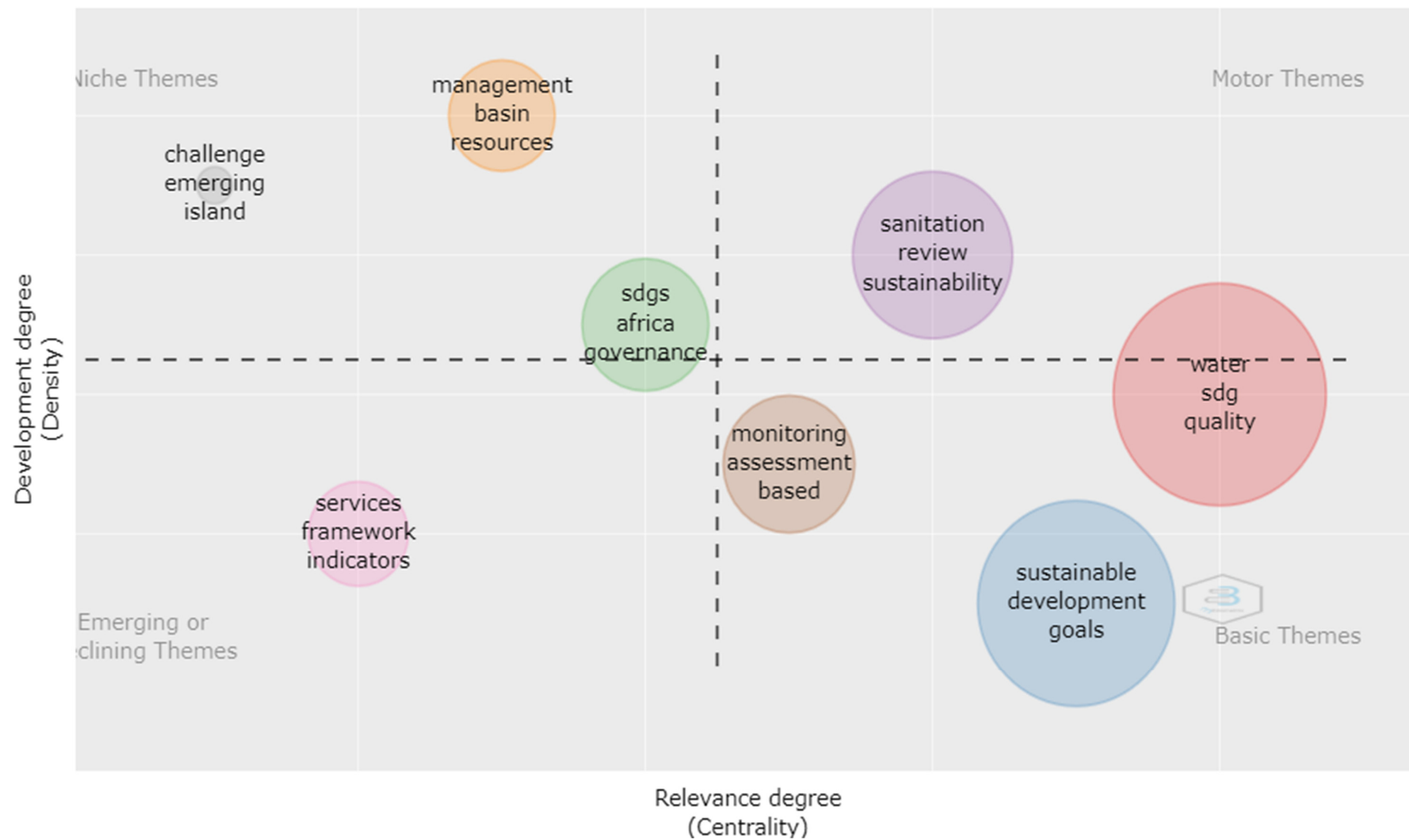
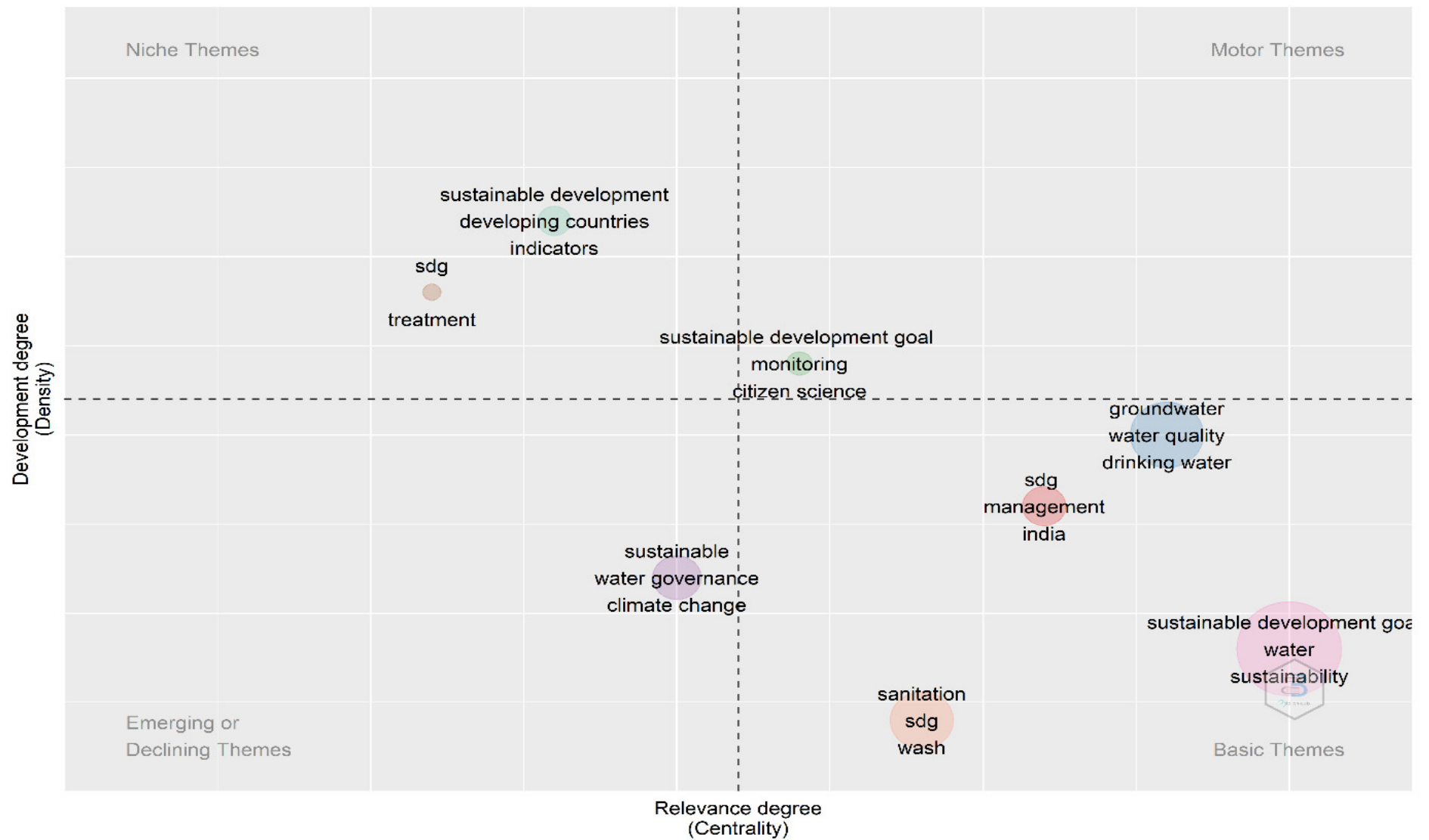**Figure S24.** Trending topics of global SDG 6 research (as per titles).

Figure S25. Thematic map of global SDG 6 research (as per Keywords Plus).

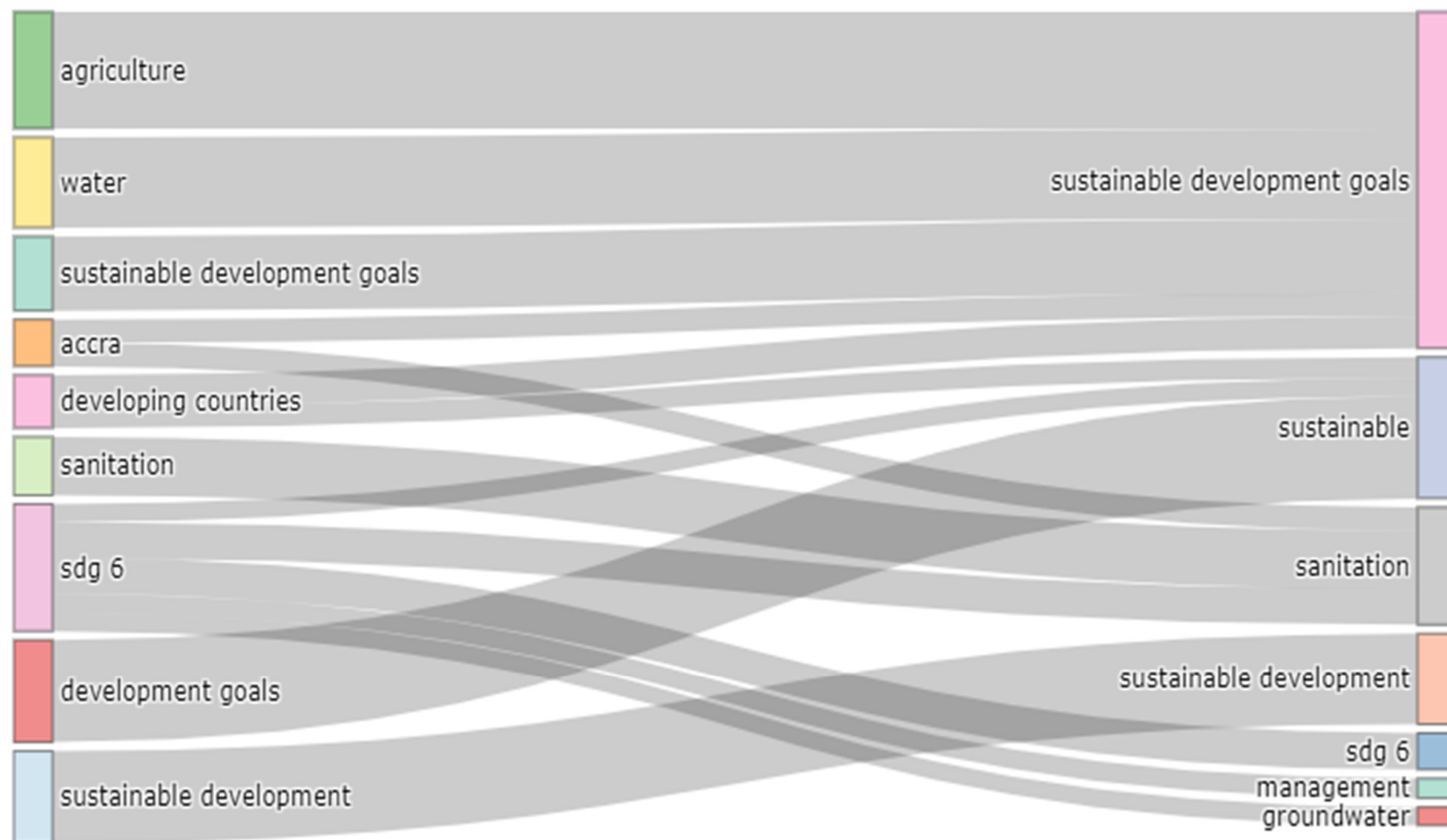**Figure S26.** Thematic map of global SDG 6 research (as per titles).

Figure S27. Thematic map of global SDG 6 research (as per content of abstracts).

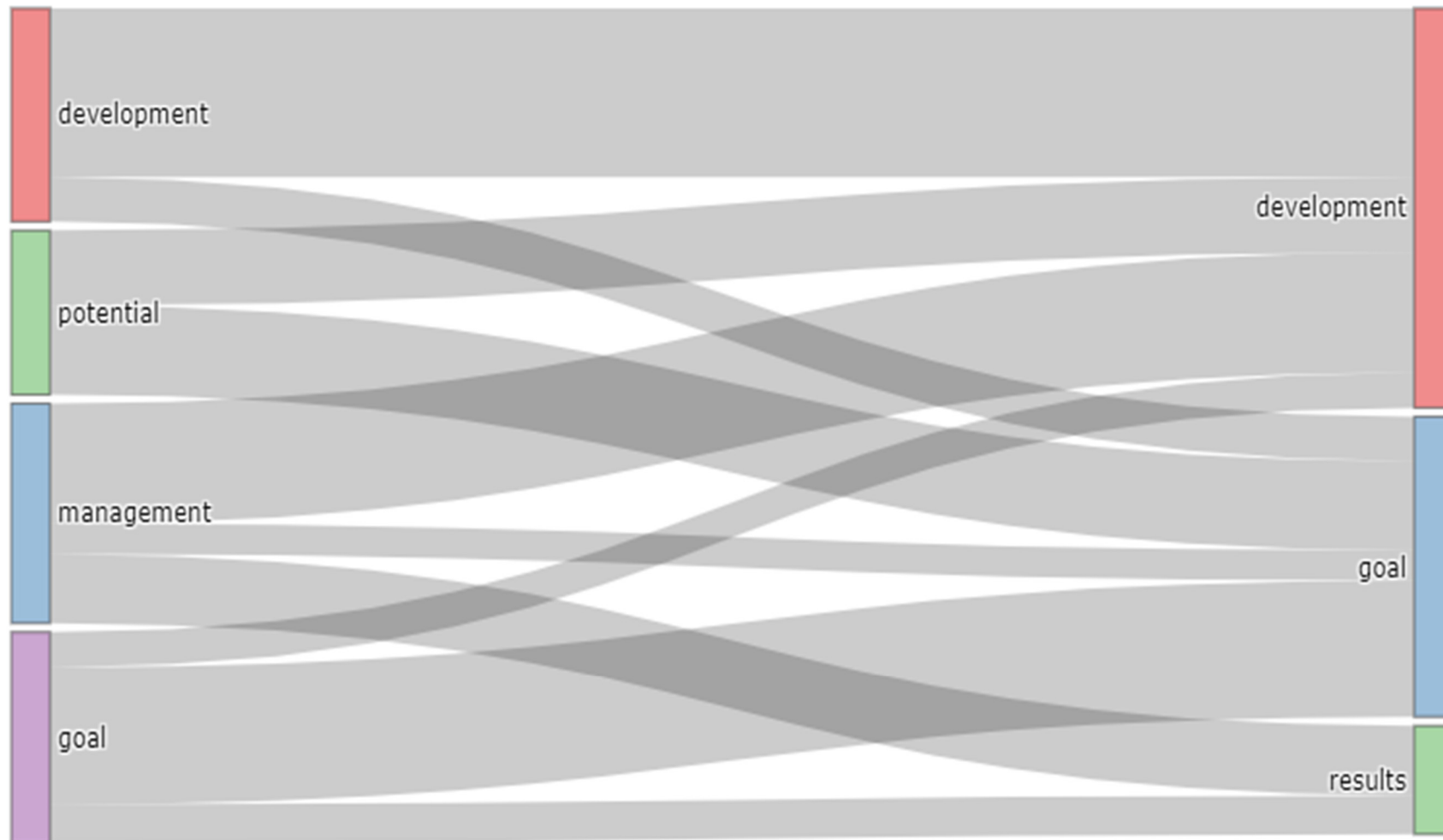**Figure S28.** Thematic map of global SDG 6 research (as per authors' keywords).

Figure S29. Thematic evolution of global SDG 6 research (as per authors' keywords).

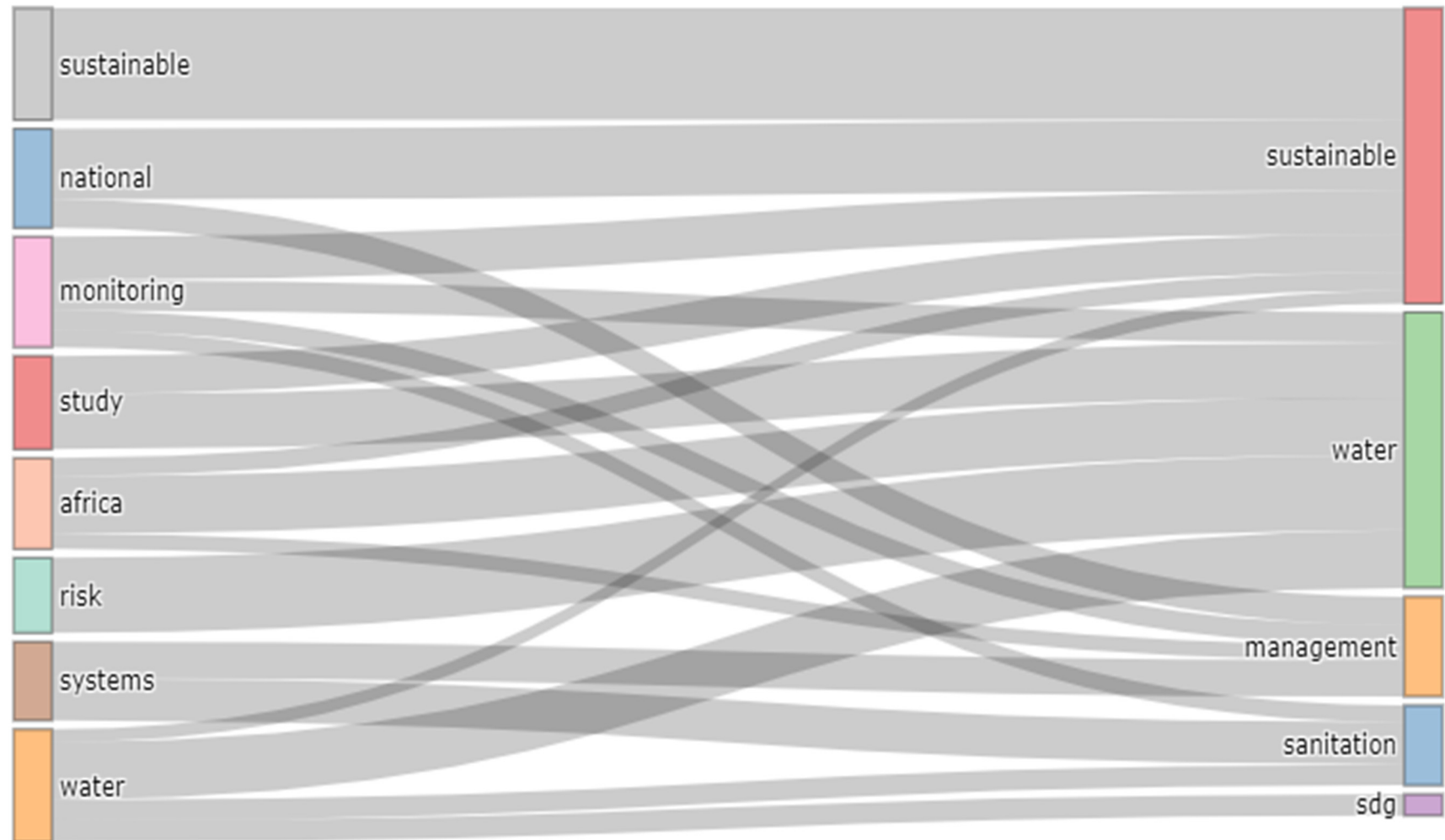**Figure S30.** Thematic evolution of global SDG 6 research (as per content of abstracts).

Figure S31. Thematic evolution of global SDG 6 research (as per titles).