

Global Research Collaboration for Vapour Intrusion

Jeroen Provoost¹ & Karen Victor²

¹ Independent researcher, Finland ORCID ID: 0000-0001-5579-2189

² Independent researcher, Finland ORCID ID: 0000-0002-3030-0229

Correspondence: Jeroen Provoost, Independent researcher, Finland. E-mail: jprovoost-research@yahoo.co.uk

Received: March 31, 2021

Accepted: April 30, 2021

Online Published: April 30, 2021

doi:10.5539/ep.v10n2p1

URL: <https://doi.org/10.5539/ep.v10n2p1>

Abstract

The complexity of the vapour intrusion (VI) transport pathway has received an ever-increased interest worldwide, and an improved and consolidated understanding of the VI issue requires collaboration between international research groups. This study uses the social network analysis methodology, applied to bibliometric authorship for VI research, to discover trends in collaboration, identify lead scientists, organisations, and countries. Furthermore, some of the external factors influencing the collaboration and productivity were assessed. The data suggests that the global research network for VI produced over a time span of 54 years 566 publications via 157 sources. The research network is composed of 437 organisations and 1053 authors from 33 countries. This suggests an increasingly active international collaborative research effort. However, inter-continental cooperation is much less than continental. The top five most central countries in the network are the USA, followed by Canada, China, The Netherlands, and Italy. The researchers with the most publications are from these five countries as well as the top organisations. The social network analysis conducted shows a good approximation of the collaborative structure for the key countries, organisations and researchers involved. Since 2010, the research community has become more stable.

Keywords: bibliometric, social, network, analysis, authorship, vapour, intrusion, Knime, Palladio

1. Introduction

Contaminated land management (CLM) is a serious issue for many industrialised countries. For example, in the European Union (EU) alone, an estimated number of 1 170 000 sites are potentially contaminated, with 342 000 confirmed, amounting to a total estimated cost of over € 5 billion, if an estimated average cost of € 10 per capita is used (JRC, 2014; Eurostat, 2018). The number of contaminated sites in the United States of America (USA) is anywhere from 300 - 400 000 sites at an estimated clean-up cost of \$ 500 billion to \$ 1 trillion (National Research Council, 1997). Site investigation and remediation make up the bulk of the costs needed to reduce the adverse human and ecological risk to an acceptable level (McAlary et al., 2011). In the management of contaminated land, volatile organic compounds (VOC) are commonly assessing for the risk of vapour intrusion (VI). VI is the process where VOC migrates from the source area, for example, the groundwater, through the soil into the foundation, and eventually to the indoor air of a building. If VOC of a hazardous substance enters the indoor air, humans can be exposed to a level that exceeds health criteria and hence poses a risk (McAlary et al., 2011).

The complexity of the VI transport pathway has received an ever-increased interest worldwide and an improved and consolidated understanding of the VI issue requires collaboration between international research groups. Networks in which scientists collaborate on research are the trademark of modern science. Scientists increasingly operate collectively via multidisciplinary approaches by which they jointly generate knowledge and look for solutions (Sonnenwald, 2007). Collaborative networking stimulates communication, the sharing of ideas and information, and can increase the productivity and generation of knowledge (The Royal Society, 2011). Complex problems, such as VI, increasingly depend on interdisciplinary research, as it involves multiple stakeholders with knowledge on the behaviour of contaminants in the soil and the building, as well as the human health impact.

Recent studies show that social network analysis (SNA) (Wasserman & Faust, 1994; Haythornthwaite, 1996) was applied to, for example, assess the cross-disciplinary approach in research programmes (Yang & Heo, 2014) or support the organizational competitive intelligence (Alcará et al., 2006).

This article uses the SNA methodology, and applies it to bibliometric authorship for VI research, to uncover collaboration trends, identify lead scientists, organizations, and countries, as to explain the impact of external

factors in scientific collaboration and productivity (Katz & Martin, 1997; Beaver, 2001; Melin & Persson, 1996; Glanzel & Schubert, 2005; Newman, 2004; Laudel, 2002; Fonseca et al., 2016).

2. Material and Method

2.1 Social Network Analysis

A social network contains members (nodes) with related relationships (edges). The presents or absence of edges between the nodes characterises the network and is investigated and quantified in an SNA. SNA allows for the identification of the most important nodes in the network, how they are grouped, and the extent of the cooperation. Nodes can be individual researchers, groups of researchers (departments), organisations or countries where researchers are located. Depending on the nodes the edges have a different meaning. For example, the edges between members, like organisations, can indicate the number of shared published work, or cooperation between countries. In an SNA the edges between nodes are related to the network and are not attributes of nodes (Fonseca et al., 2016). Scientists collaborate to discover knowledge that often requires different skills for complex problems (Sonnenwald, 2007; Katz & Martin, 1997). Published work, such as peer-reviewed papers or conference proceedings, reflects the collaboration between scientists via authorship (Newman, 2004). Authorship analysis reveals the cooperation patterns between scientists, the organisations they work for and the countries in which they are situated (Melin & Persson, 1996; Glanzel & Schubert, 2005; Newman, 2004). The steps applied for this authorship analysis are collection, normalisation, and curation of publications, followed by the calculation of metrics and visualisation.

2.2 Collection, Normalisation and Curation of Author Information

Published work was obtained from several sources:

- Science Direct (<https://www.sciencedirect.com/>)
- Web of Science (<http://apps.webofknowledge.com>)
- Google scholar (<https://scholar.google.com/>)
- Research Gate (<https://www.researchgate.net/>)

The search terms used were vapor or vapour and intrusion in the title or abstract. The resulting database covers the following type of publications: papers from journals, chapters from books, Bachelor, Master or Doctoral theses, articles in bulletins or magazines, research reports mostly from project work, and conference proceedings. Journals, chapters in books and theses are considered to be peer-reviewed, while the peer-review process for articles, reports, and conference proceedings is not clearly defined.

From each publication, the letters of the first name and full last name of each author were recorded in a spreadsheet, checked for accuracy, and recorded together with the organisations name, address, and geographical coordinates (longitude and latitude). As each authors' name and organisation are core components of the SNA, considerable time was spent verifying the accuracy of the data collected by manually checking spelling errors, omissions, the similarity of names and name changes. For authors who reported multiple organisations the first indicated organisation was selected. The department or faculty of the authors was not recorded, for example, some researchers work for the same university, but in different departments. For some publications more than one corresponding author is proposed. In this case the first was selected. Authors of reports were recorded individually, and if not presented in the report, the publishing organisation was recorded. Some of the organisations do no longer exist, so the geographical location was put in the centre of the city. Special characters in names and organisations were simplified to base letters (ü becomes u). Names of Journals are written in full (not abbreviation). Publications with many authors, like for some of the reports, were investigated and reviewers were not considered while contributing authors were retained (Adams, 2012). The standardizing, curation, and profiling of data was performed with the Knime data analysis platform (Berthold et al., 2009; Tiwari & Sekhar, 2007). The aim was to consolidate names and organisations to ensure the acknowledgement of the cooperation because inaccurate data can generate errors in the presentation of the network (Wang et al., 2012; Barbastefano et al., 2013).

2.3 Calculation of Metrics and Visualisation

The authors, organisations, and journals were then transformed into matrices that provide the nodes (like author or organisation) and their relationship via edges (like the shared publications or country). The edges indicate the intensity of the cooperation (nodes share the authorship of a publication). The authorship cooperation is considered to be reciprocal, so non-directional. A network provides insight into the strategic importance of a node relative to other nodes, and to identify the cooperation's centrality measures are calculated (Freeman, 1978; Ramadan, Alinsaif & Hassan, 2016). Several measures were calculated to understand the cooperation in the network. The

below topological measures were used.

2.3.1 Degree-Based Measurements

The degree measures provide insight into the number of neighbours a node has. The more connections a node has the more important the node is in the network. A higher centralisation indicates that one or more nodes are connecting with many others, a low centralisation indicates that the connections between nodes are more evenly distributed.

The following degree centralities are calculated (description from Knime, 2021):

- 1) The node degree counts the number of incident edges. The number of nodes and links represents the network size. The percent degree is the percentage of edges in comparison to the total amount of edges.
- 2) The in- and out-degree count the number of incoming and outgoing edges. The percent incoming and outgoing degree is the percentage of edges in comparison to the total amount of edges.

2.3.2 Shortest-Path-Based Measurements

Distance measures were used to express the cooperation between nodes in the network. The smaller the distance between the nodes, the more related any the nodes are (Gong et al., 2016).

The following distance-based measures were calculated (description from Knime, 2021):

- 3) The closeness centrality divides the number of nodes of the component by the sum of all distances from the analysed node to all other nodes within the component (Knime, 2021). The node with the highest value is the most central node of its component. The closeness centrality (Sabidussi, 1966; Freeman, 1978) measures the average distance of the shortest path from a node to other nodes or is the closest to all other nodes. Centralization refers to the degree to which edges are concentrated in a few nodes in the network and evaluates the presents of dominant nodes in the network.
- 4) The node weight sum is the sum of the weight of all incident edges, and the average node weight (del Rio, Koschützki & Coello, 2009).
- 5) The clustering coefficient analyses the neighbourhood of the node. If they form a clique the coefficient is 1, and if no neighbour is connected to another neighbour it is 0.
- 6) The Barycenter assigns scores to each node according to the sum of its distances to all other nodes. The edge weight represents similarity, and this is converted to a distance by using the JUNG Framework (O'Madadhain et al., 2005). It indicates what node exerts the most influence in the network (Viswanath, 2009; Gadat, Gavra & Risser, 2016).

2.3.3 Importance Measures

- 7) The Hubs and Authority centrality assigns hub and authority scores to each node (León, 2013; Xutao, Michael & Yunming, 2012; Kleinberg, 1999). Hub and Authority centrality focus on the structure of the network and determine its importance according to their positions on a graph (Marra et al., 2015) and is affected by the total relations that occur outside the node (Farhan, Darwiyanto & Asror, 2019). The node is a hub when it has edges to authoritative nodes. It is an authority if it is referenced by 'hub' nodes. This implies a mutually reinforcing centrality. Therefore, a high hub node points to many good authorities, and a high authority node receives from many good hubs. To calculate the score the JUNG framework is used (O'Madadhain et al., 2005).

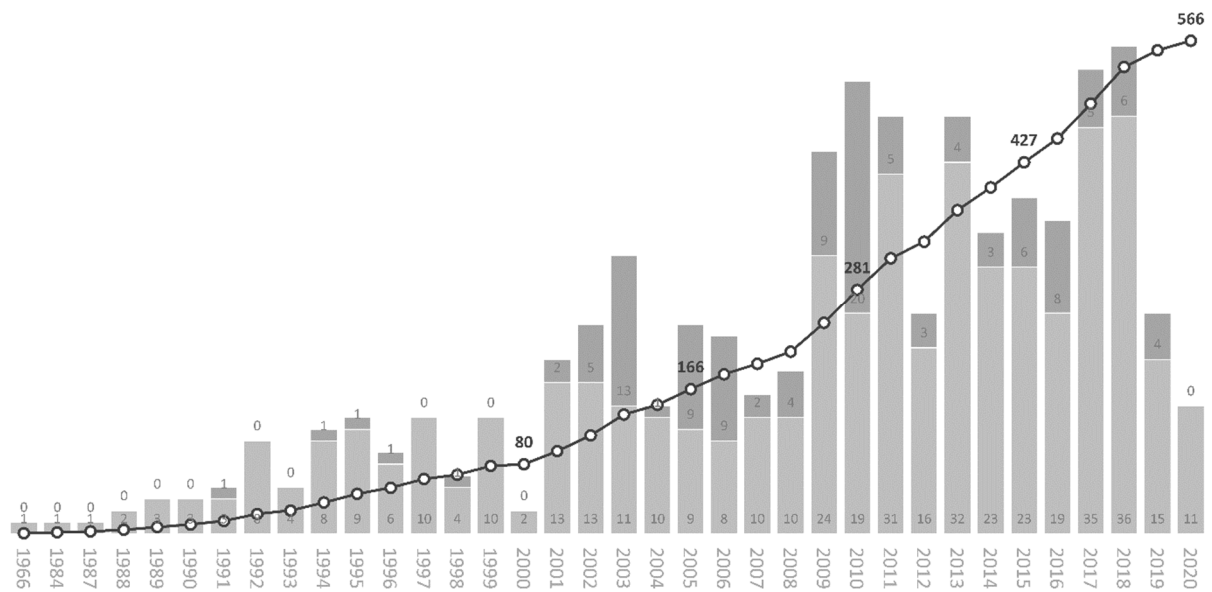
Palladio (Edelstein, Coleman & Findlen, 2017) was applied for the geographical visualization of the cooperation, while Knime (Berthold et al., 2009) was used to visualise the various networks (Fillbrunn et al., 2017). Network drawings were generated by using the Kamada-Kawai algorithm (Kamada & Kawai, 1989). A network describes the overall structure with the main nodes (authors) and their different connections (edges) (Valente, 2010), and factors that may have affected its configuration could be assessed against the network.

A workflow was built in Knime to produce tables that provide an overview of the different type of publications by year, the productivity of authors, organisations, countries, journals, as well as their relationships in the network (Tsay & Shu, 2011; Rosas et al., 2011; Ellegaard & Wallin, 2015; Merigó & Yang, 2017; Newman, 2004).

3. Results

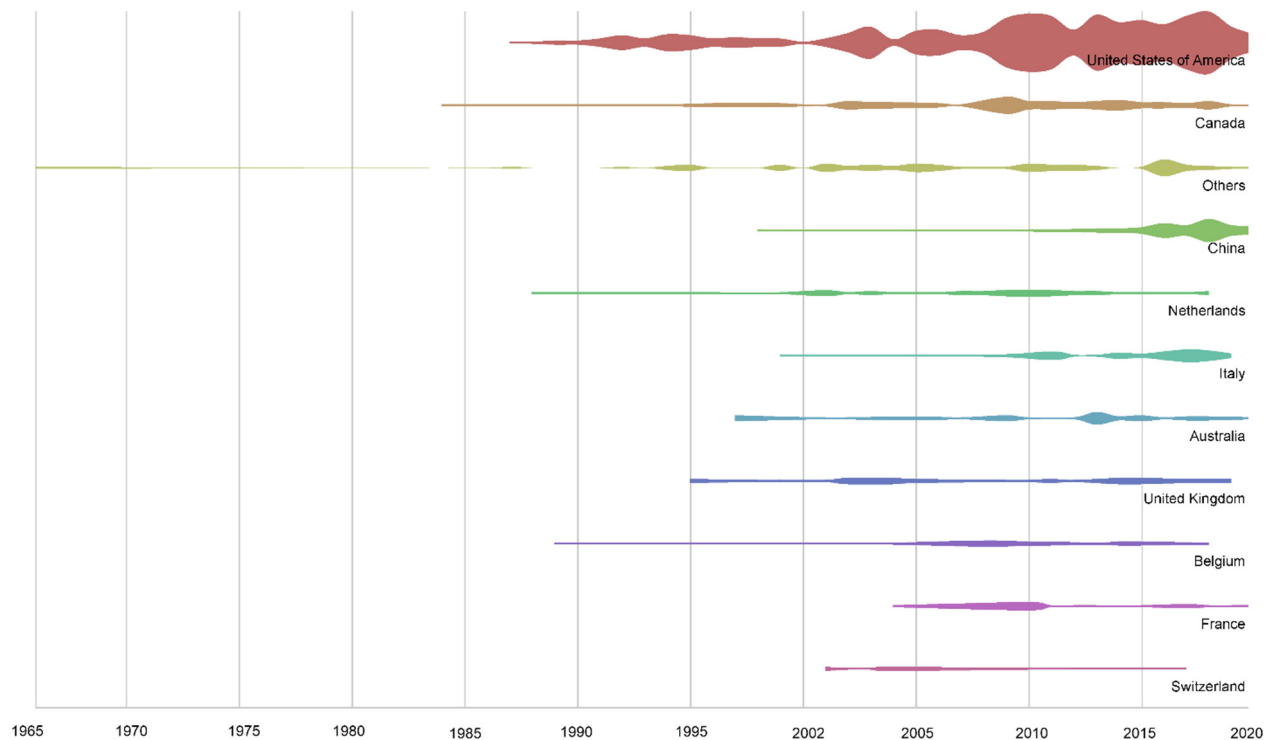
The below results reflect the international collaboration on research efforts for VI. Tables and figures are used for descriptive findings, overall performance, and research output. The results identify and quantity productivity and top impact, as well as collaboration patterns.

3.1 Productivity



a: number of publications by type over time.

—○— represents the cumulative number of publications, ■ the peer reviewed publications like papers, book chapters and theses, ■ the other publications like reports, conference proceedings and articles.



b: number of publications by country over time

Figure 1. Number of publications by year versus type (a) and country (b)

The search terms resulted in a total of 566 publications, spanning 54 years (1966 till half 2020). Figure 1a shows that 78% of the publications are from peer-reviewed sources, like journal papers, book chapters, and theses, while

29% are from other sources, like reports, conference proceedings, or articles in magazines. Just over 70% of the publications originate from scientific journals. Between 1966 and 2009 (43 years) around 50% of the work on VI was published, and between 2010 and half of 2020 (± 10 years) the rest. In the last 5 years alone 25% of work was published. The area graph (figure 1b) shows that since 1988 the USA published an increasing amount of work on VI, while other countries start contributing since 2010.

Table 1 to 4 show the productivity (top 15) of respective countries, journals, authors, and organisations for all published work.

Table 1. Productivity of countries

Country	# Publications	# Publication sources	# Authors	# Organisations
United States of America	372	107	566	227
Canada	66	30	78	29
China	41	19	74	20
Netherlands	33	17	51	15
Italy	30	17	25	10
Australia	29	18	37	20
United Kingdom	28	16	47	28
Belgium	17	14	17	4
France	16	13	41	22
Switzerland	10	4	17	5
Denmark	8	5	16	7
Germany	7	7	21	7
South Korea	6	4	14	8
Sweden	5	3	7	4
Brazil	5	4	9	6

A total of 33 countries published work on VI, reflecting an international research network. It is evident that the USA generates the most publications, and so has most authors and related organisations involved in a wide variety of published sources. Table 1 reveals that the USA is the most productive when it comes to generating knowledge on the topic of VI, followed by Canada, China, Netherlands, Italy, and Australia.

Table 2. Productivity of journals

Name journal	# Publications	# Authors	# Organisations	# Countries
Soil and Sediment Contamination: An International Journal	115	266	120	22
Environmental Science and Technology	43	111	47	10
Ground Water Monitoring and Remediation	21	78	42	9
Journal of Contaminant Hydrology	16	44	22	9
Journal of Hazardous Materials	16	49	21	9
Remediation	15	36	22	5
Science of the Total Environment	15	66	26	7
Water Resources Research	11	20	12	4
Environmental Science Processes and Impacts	9	28	14	6
Building and Environment	8	20	9	3
Journal of Soil Contamination	7	14	10	3
Atmospheric Environment	6	18	8	3
Vadose Zone Journal	5	25	11	4
Human and Ecological Risk Assessment: An International Journal	4	11	5	2
Journal of Environmental Quality	4	16	9	3

The publications were published by 157 different sources, of which 90 (57%) were scientific peer-reviewed journals. The most productive journal on the topic of VI, with the most authors, affinities and, countries, is the journal of Soil and Sediment Contamination: An International Journal, followed by Environmental Science and Technology, Ground Water Monitoring and Remediation, Journal of Contaminant Hydrology, Journal of Hazardous Materials, Remediation, and Science of the Total Environment.

Table 3. Productivity of authors

Author	# Publications	# Journals	# Books	# Bulletins /	# Conferences
Suuberg EM	39	14	2		1
Yao Y	33	13	1		1
Pennell KG	29	15	2		1
Johnson PC	33	6	1	1	2
McAlary TA	25	5	2	1	5
McHugh TE	23	8	1		2
Verginelli I	18	9			1
Provoost J	16	6	3		3
Shen R	15	10	1		
Hers I	18	6			1
Schumacher BA	13	5			2
Swartjes FA	13	6	1		
Hartman B	15	2	1	1	
Truesdale RS	13	4		1	1
Gorder KA	13	4	1		

The publications were written by 1053 authors. The authors who were most productive and published in a variety of journals are Suuberg EM and Yao Y, followed by Pennell KG, Johnson PC, McAlary TA, McHugh TE, and Verginelli I. A geometric mean of almost 3 authors per publication was observed.

Table 4. Productivity of organisations

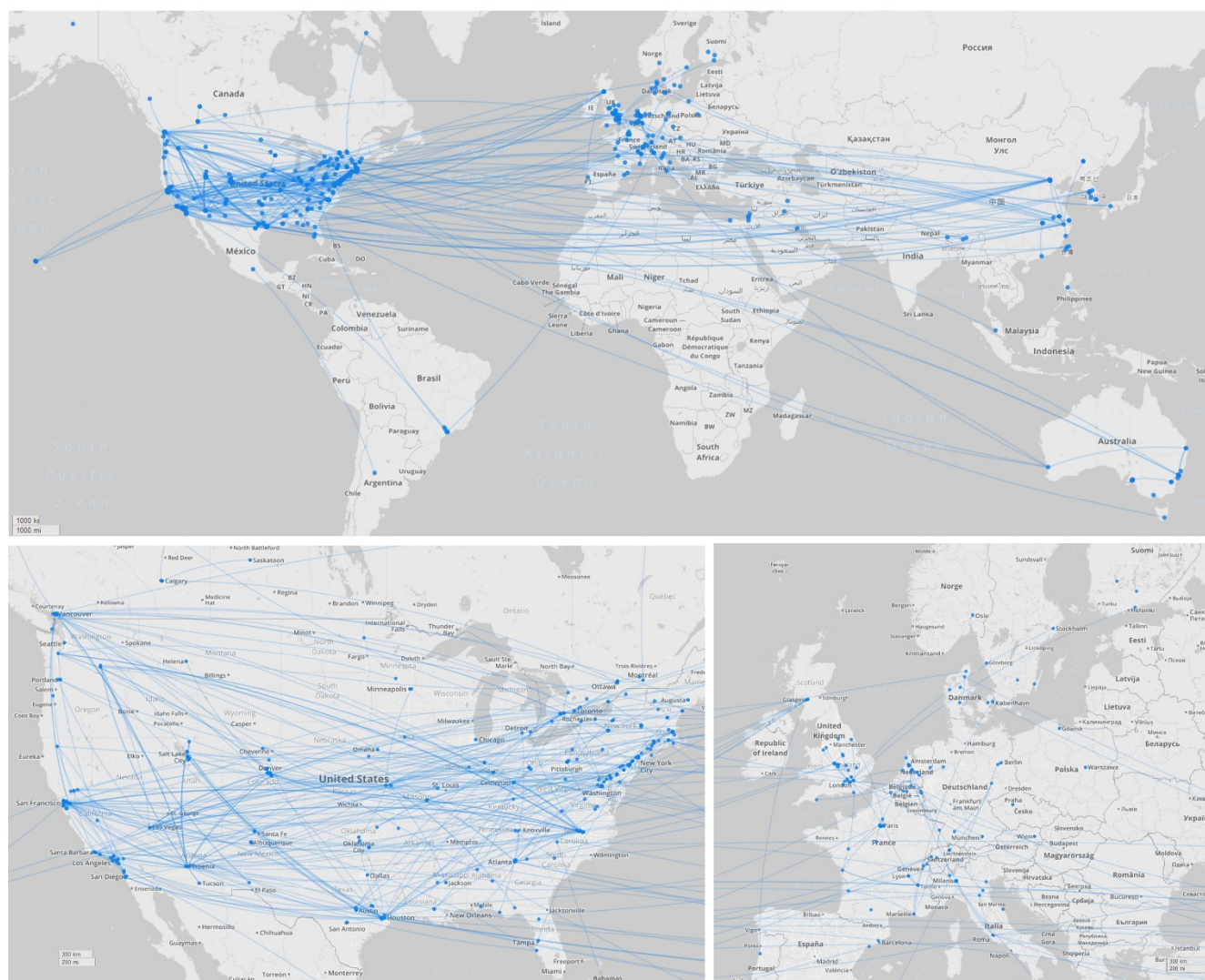
Name organisation	# Publications	# Authors	# Countries	# Journals	# Books	# Bulletins / magazines	# Conferences
Environmental Protection Agency	42	28	1	5	2	3	5
Brown University	40	9	1	14	2		1
Arizona State University	33	10	1	7		2	2
Geosyntec Consultants Inc.	31	18	2	6	2	2	6
Groundwater Services Inc.	25	17	1	8	1		2
Golder Associates Ltd.	18	12	3	7			1
National Institute of Public Health and the Environment	18	18	1	6	1		
Shell Global Solutions	18	5	1	6	1	1	
University of Rome "Tor Vergata"	18	5	1	9			1
Zhejiang University	18	18	1	9			
U.S. Air Force	16	5	1	5	1		1
Flemish Institute for Technological Research	15	10	1	6	3		3
University of California	14	26	1	8			
Hartman Environmental Geoscience	13	1	1	1	1	1	
RTI International	13	4	1	4		1	1

A total of 437 organisations were involved, of which 143 (33%) were universities or schools. The most productive organisation is the (USA) Environmental Protection Agency, closely followed by Brown University, Arizona State University, and Geosyntec Consultants Inc.

The organisation with the most authors involved are the Environmental Protection Agency and the University of California, followed by Geosyntec Consultants Inc., National Institute of Public Health and the Environment, Zhejiang University and Groundwater Services Inc. Brown University published most papers in scientific journals.

3.2 Collaboration

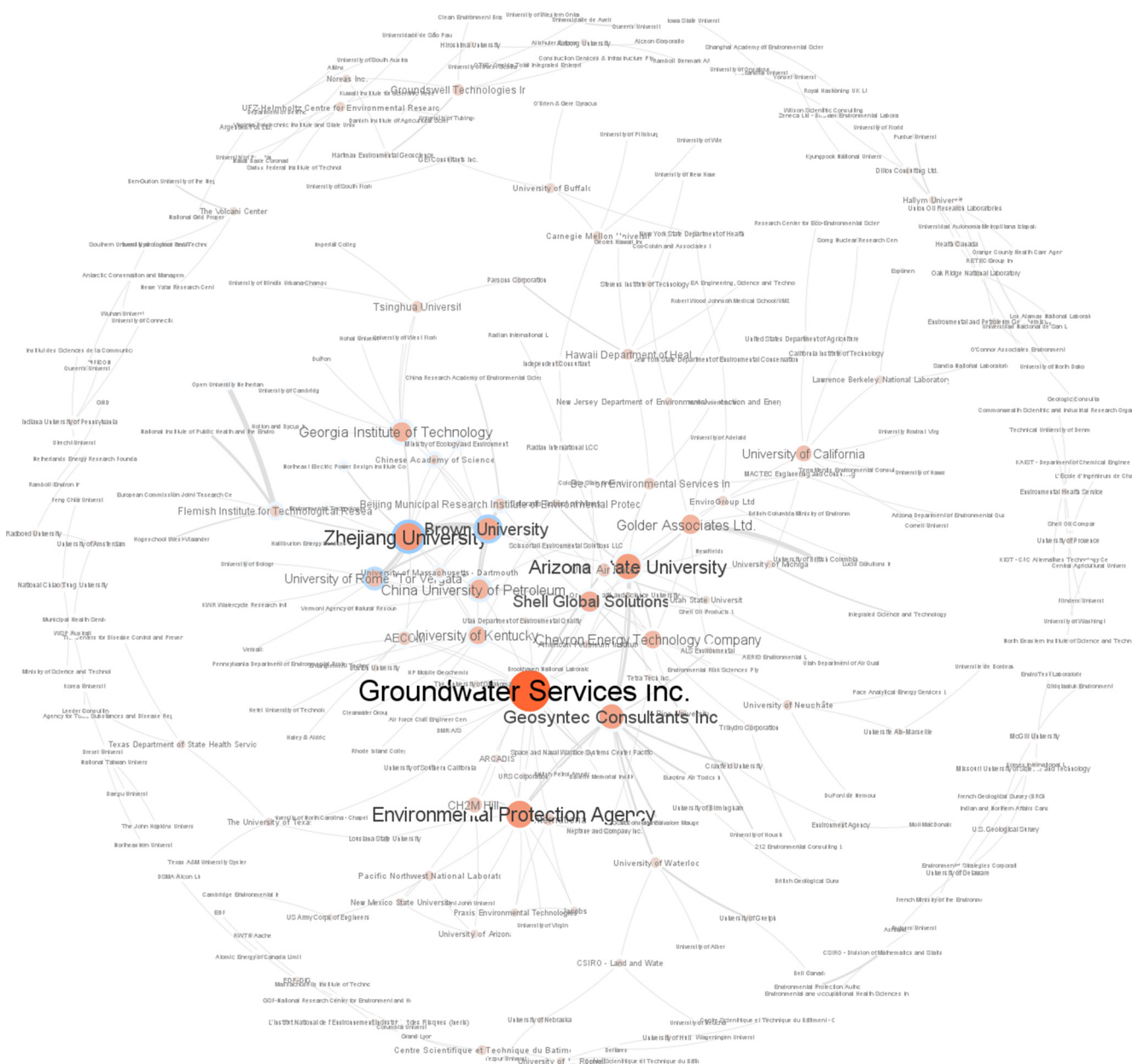
The below figures 2 to 5 display the scientific cooperation between organisations, authors, and countries. To visualise the cooperation between organisations and countries the number of publications was aggregated.



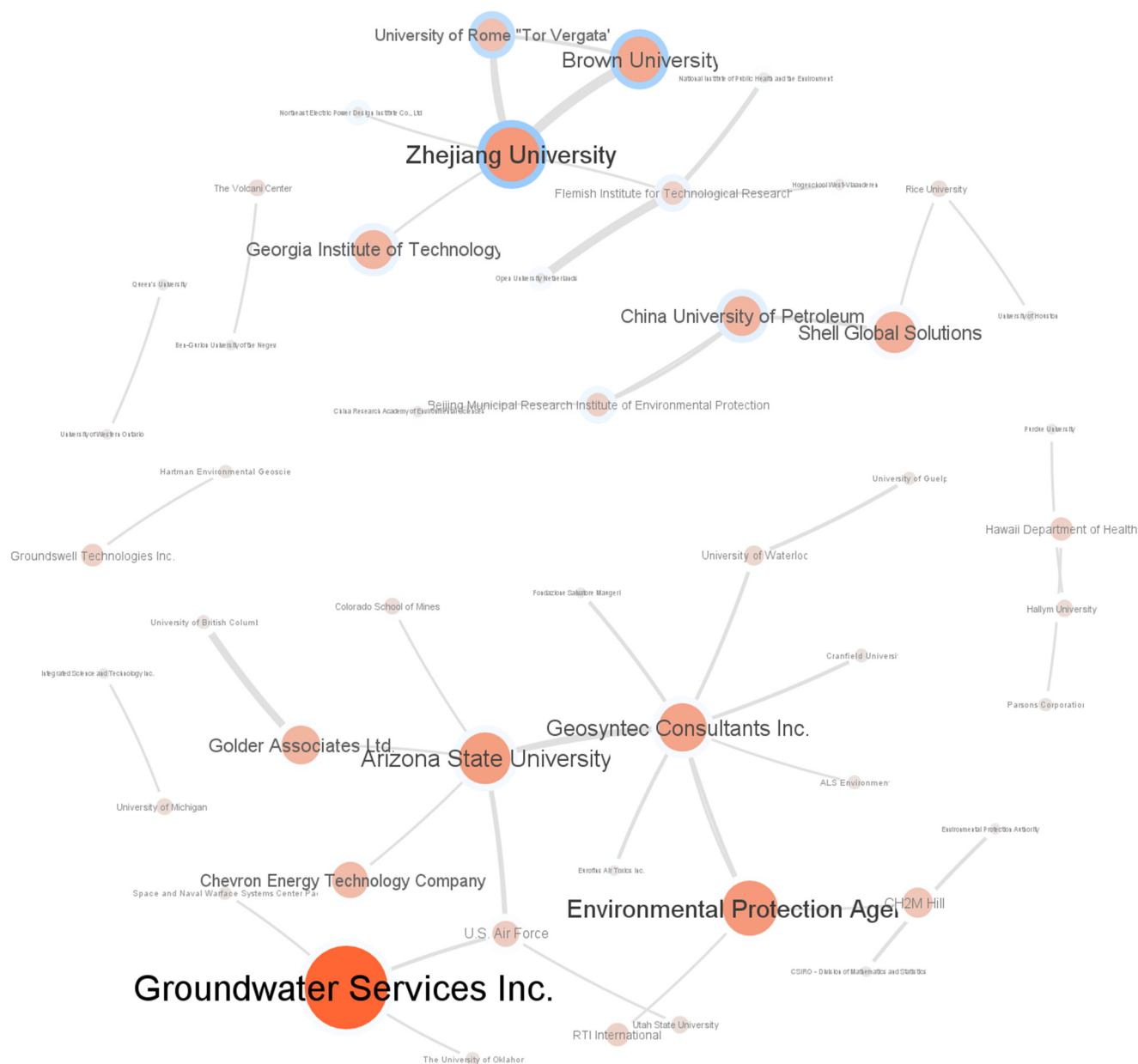
● represent the organisation geographical location based on the longitude and latitude coordinates, — represent the cooperation as a result of a joint publication as author. The lower two images provide an enlargement for the USA plus Canada, as well as for Europe.

Figure 2. Geographical representation of the cooperation between organisations

Figure 2 shows that most organisations are situated in the USA and Canada. Cooperation between European or Asian countries is limited. Inter-continental cooperation occurs, mostly between the USA and Canada, and both countries cooperating with affiliates in the EU or Asia.



a: cooperation between all organisations



b: cooperation between organisations that collaborated more than once (filtered version of a)

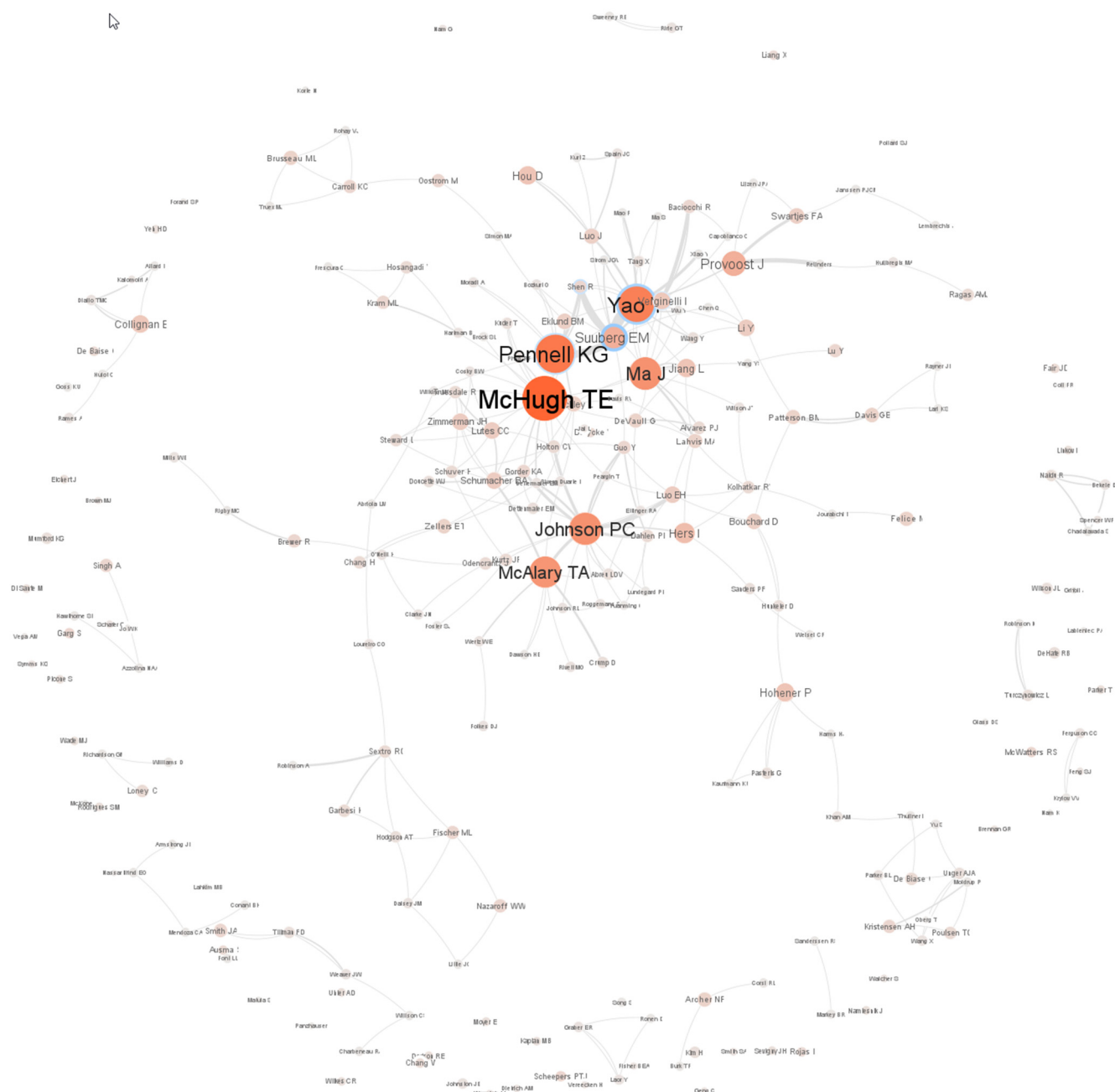
Figure 3. Visualisation of the cooperation between organisations (a. all organisations, b. filtered)

Figure 3a/b depicts the node degree by utilising the circle size and intensity of the red colour. So, a larger and darker red circle indicates more incident edges (cooperation). The blue outline stroke width represents the authority centrality. The darker blue the outline stroke the more authority the organisation has in the network. The width of the edges indicates the strength of the collaboration (publishing more jointly).

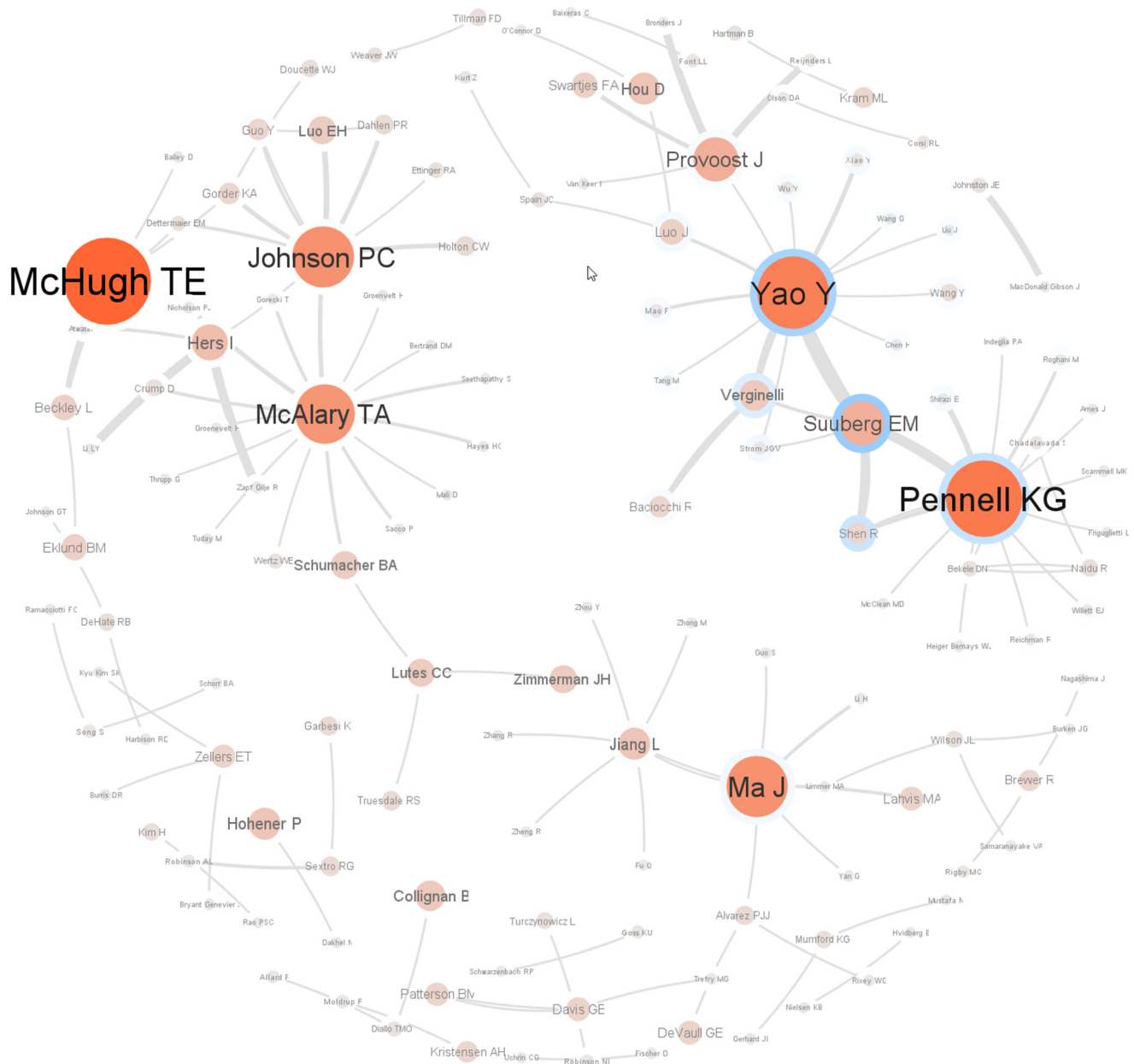
Figure 3a reveals clusters of cooperating organisations, with some organisations (nodes) being central in the network and having established a strong cooperation (many edges). Some of the central organisations have a high authority. Authors from the organisation Groundwater Services Inc. cooperate most with others, followed by Zhejiang University, Environmental Protection Agency, Arizona State University, Geosyntec Consultants Inc., Brown University, China University of Petroleum, Georgia Institute of Technology, and Golder Associates Ltd.

Figure 3b provides a filtered view of figure 3a where the edge is above 1, meaning that organisation collaborated more than once, therefore highlighting the key organisations in the network. The organisations with a high node degree (cooperation) are not necessarily those with authority in the network. The organisations with the highest

authority scores are in descending order the Zhejiang University, Brown University, University of Rome ‘Tor Vergata’ and the China University of Petroleum. The figure reveals that Groundwater Services Inc. does not frequently publish work with organisations that are themselves authorities. The Zhejiang University, Brown University, and the University of Rome ‘Tor Vergata’ are research organisations that collaborate as authorities in the field of VI, given the width of the edges. The Zhejiang University and the University of Rome ‘Tor Vergata’ have a well-established cooperation. Other clusters of collaborations can be observed between Geosyntec Consultants Inc., a consultancy firm, the US Environmental Protection Agency, and Arizona State University. The width of the edge indicates that they frequently cooperated. Golder Associates Ltd., a consultancy firm, has an established cooperation with the University of British Columbia.



a: cooperation between all authors



b: cooperation between authors that collaborated more than once (filtered version of a)

Figure 4. Visualisation of the cooperation between authors (a. all authors, b. filtered)

Figure 4a/b shows the same measures as figure 3a/b, but nodes represent authors of publications. The network (figure 4a) shows that many cooperations are in the periphery of the network, suggesting that the cooperation is not sustained. However, central in the network several authors cooperate frequently, like McHugh TE, Pennell KG, Yao Y, Johnson PC, McAlary TA, and Ma J. They are well connected but do not exhibit authority in the network (darker blue outline stroke). Figure 4b is a filtered version of the full network, and highlights the more established cooperation between authors, like McHugh TE, Hers I, Johnson PC and McAlary TA, or the connection between Ma J, Jiang L, and Lahvis MA. The figure suggests established cooperations between Suuberg EM, Yao Y, Verginelli I, Shen R and, Penell KG, that contains authors with authority in the field of VI.

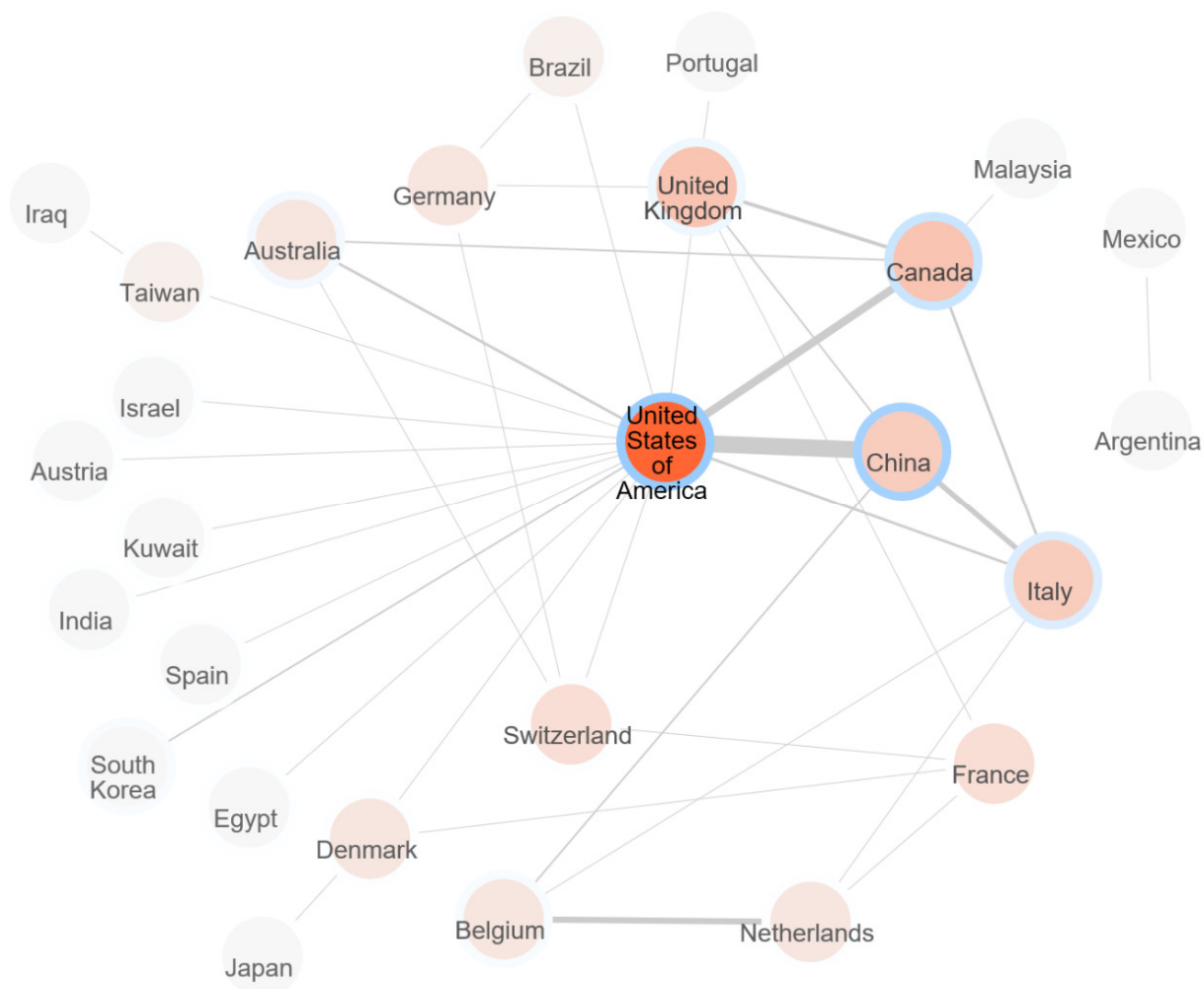


Figure 5. Visualisation of the cooperation between countries

The underlying research networks between organisations and researchers are reflected in the network of countries. Figure 5 shows the collaboration between countries. A darker red node (higher node degree) indicates that the country cooperates with other countries, while the blue outline stroke reveals the authority of a country in the network. The figure reveals that the USA collaborates most with other countries, specifically with China, and Canada, given the width of the edges. Belgium and the Netherlands also collaborate frequently on the topic of VI. The USA, China, Canada and Italy are considered authorities in the country network and collaboration.

4. Discussion and Conclusions

This study investigates scientific bibliometric authorship for VI research to assess collaboration trends between (identified lead) scientists, organisations, and countries. External factors influence research collaboration and scientific productivity and are to the extent possible addressed.

The data suggests that the global research network for VI produced over a period of 54 years, 566 publications (figure 1) through 157 sources (table 2). Three-quarter of the publications are made available from sources that apply a peer review process, and in the last 10 years around 50% of the work on VI was published (figure 1). The research network on VI is composed of 437 organisations (table 4) and 1053 authors (table 3) from 33 countries (table 1). This suggests increasing active international collaborative research effort for VI. However, the inter-continental cooperation (e.g. USA – China or Canada – United Kingdom) is much less than continental (e.g. USA – Canada or Belgium - Netherlands) (figure 2). A factor affecting cooperation is the importance given to the exposure route of VI in the various legislative frameworks for CLM. Since the Love Canal Tragedy in 1978 (Phillips, Hung & Bosela, 2007) and BKK landfill in West Covina southern California (Wood & Porter, 1987) VI start receiving attention in the context of CLM in the USA, resulting in the Superfund programme. Around ten

years later European countries started regulating CLM. Norway enacted the Pollution Control Act of 1981, Greece the Environmental Law was enacted in 1986, and The Netherlands voting the Soil Protection Act into force in 1987 (Ferguson, 1999). Likely regulating CLM is linked to funding made available for VI research. As a result, the top five most central countries are the USA, followed by Canada, China, Netherlands, and Italy (table 1). The researchers with the most publications are from these five countries (table 3) as well as the top organisations (table 4). The same applies to the centrality measures (figure 3a/b, 4a/b, and 5).

Measures, like the degree centrality, are a proxy for cooperation in a network but do not necessarily represent the volume of published work. For example, the United Kingdom is amongst the countries with the highest node degree (figure 5) but is behind the Netherlands and Italy in the number of publications (table 1). Despite Canada publishing more papers than China (respectively 66 and 14 publications) (table 1) it ranks lower in degree centrality (figure 5), which suggests a research focus that is somewhat more nationally oriented (Fonseca et al 2016). The presence of the USA and China centrally in the network reflects a commitment to the issue of VI research.

Measures of the organisations (figure 3a) and authors (figure 4a) network show low density and centralization values associated with many peripheral communities (Fonseca et al 2016) which suggest sparse cooperation patterns among organisations and authors (figure 3b and 4b), indicating opportunities for increasing the cooperation. Some organisations do not rank high for the degree centrality but fulfil the role of intermediate linking groups together. The Flemish Institute for Technological Research, Tsinghua University, and Hawaii Department of Health are such examples (figure 3a). This provides them with some influence and control of information.

The SNA conducted shows a good representation of the network structure and key countries, organisations and researchers involved and provides suggestions for further research. Since 2010, the research community has become more stable. Further temporal analysis could provide insights into the evolution and changes of the research network. Adding additional data can add to the perspective on VI research, for example adding authors from literature references or data collection to determine what drives researchers and how they maintain their network (Laudel, 2002).

Although evidence of successful SNA based policies is limited, some examples highlight the potential use (Morel et al., 2009; Bender et al., 2015; Eslami, Ebadi & Schiffauerova, 2013; Lander, 2013). The results from this SNA shows:

- 1.) Fragmentation of the network (figure 3a/4a) and so a need to consolidate cooperation's amongst researchers, organisations, and countries. The countries with most published work and collaborations are from high-income economies (figure 5), ergo the need arises to cooperate with middle- and low-income countries as to transfer knowledge on CLM and the VI pathway.
- 2.) Few organisations are central in the network and consolidate knowledge (figure 3a/b). Universities created a network in which they cooperate, public bodies (Environmental Protection Agency, Beijing Municipal Institute of Environmental Protection) and in the USA also with consultancy companies. The latter is prominent in the network (Groundwater Services Inc., Geosyntec Consultants Inc., Golder Associates Ltd.), as well as commercial companies (Shell Global Solutions, Chevron Energy Technology Company, Groundswell Technologies Inc.). The cooperation happens mainly in the USA and Canada, less in the EU and Asia.
- 3.) Country-specific settings or legislation for CLM can influence the cooperation in the VI network. In the USA the Superfund program (EPA, 2021) is responsible for financing the clean-up of the most contaminated land and required organisations to cooperate to bring knowledge of many aspects together. The EU has funded several networks like NICOLE (Network for Industrially Co-ordinated Sustainable Land Management in Europe) and CLARINET (The Contaminated Land Rehabilitation Network for Environmental Technologies in Europe). The NICOLE network aim was to coordinate sustainable land management in the EU, and increase the cooperation between various players (academia, service providers and industry) for the development and application of sustainable technologies (NICOLE, 2002, 2012). CLARINET's primary objective was to develop technical recommendations for decision-making on the rehabilitation of contaminated land in the EU. Furthermore, to identify and report on research and development needs (CLARINET 2002; Bardos, 2003). Pertaining to the exposure path of VI the cooperation between industry and research is visible in the USA, not in the EU, despite the aim of the funded networks. Why this is not visible in the results needs further research.

SNA has proven to be a useful tool for retrieving the composition and cooperation in the VI network of researchers, organisations, and countries. However, SNA has limitations, as researchers do not only cooperate via published work. Cooperation does not imply knowledge sharing; however, authorship requires a level of cooperation beyond the exchange of information. Quantitative SNA is unable to analyse the reason and motivation for the network structure. This issue can be assessed by using qualitative SNA methods like interviews or a more in-depth data

gathering on the cooperation (Kolleck, 2013).

Conflict of interest

The authors declare that they have no conflict of interest with the topic or company that distributes the software used in this paper.

Acknowledgement

This research received no funding from any public, commercial, or non-profit organization. No organization had any role in or contribution to the content of this paper.

References

- Adams, J. (2012). Collaborations: the rise of research networks. *Nature*, 490(7420), 335–336. Retrieved from <http://go.nature.com/zzwn8z>
- Alcará, A. R., Tanzawa, E. C. L., Di Chiara, I. G., Tomaél, M. I., Junior, P. P. D. M. U., Heckler, V. C., Rodrigues, J. L., & Valente, S. S. (2006). Social networks used as a strategic instrument for competitive intelligence. *Transinformação*, 18(2), 143-153. <https://doi.org/10.1590/S0103-37862006000200006>
- Barbastefano, R. G., Souza, C., Costa, J. S., & Teixeira, P. M. (2013). Names and its impacts on social networks properties: a study in a co-authorship network on sustainability. *Perspectivas em Ciência da Informação*, 18(3), 78-95. <https://doi.org/10.1590/S1413-99362013000300006>
- Bardos, P. (2003). A review of the Contaminated Land Rehabilitation Network for Environmental Technologies in Europe (CLARINET). Part 2: Working Group findings. *Land Contamination & Reclamation*, 11(1). <https://doi.org/10.2462/09670513.616>
- Beaver, D. D. (2001). Reflections on scientific collaboration (and its study): past, present, and future. *Scientometrics*, 52(3), 365-377. <https://doi.org/10.1023/A:1014254214337>
- Bender, M. E., Edwards, S., von Philipsborn, P., Steinbeis, F., Keil, T., & Tinnemann, P. (2015). Using co-authorship networks to map and analyse global neglected tropical disease research with an affiliation to Germany. *PLoS Neglected Tropical Diseases*, 9(12). <https://doi.org/10.1371/journal.pntd.0004182>
- Berthold, M. R., Cebren, N., Dill, F., Gabriel, T. R., Kötter, T., Meinl, T., Ohl, P., Thiel, K., & Wiswedel, B. (2009). KNIME - the Konstanz information miner. *ACM SIGKDD Explorations Newsletter*, 11(1), 26. <https://dx.doi.org/10.1145/1656274.1656280>
- CLARINET. (2002). *Sustainable management of contaminated land: An overview, report*. Retrieved from https://clu-in.org/wales/download/1CLARINET_RBLM_report.pdf
- del Rio, G., Koschützki, D., & Coello, G. (2009). How to identify essential genes from molecular networks? *BMC Systems Biology*, 3(1), 102. <https://doi.org/10.1186/1752-0509-3-102>
- Edelstein, D., Coleman, N., & Findlen, P. (2017). *Networks in History: Data-driven tools for analyzing relationships across time, Final white paper*. Retrieved from <https://hcommons.org/deposits/download/hc:12386/CONTENT/hk-50087-13.pdf/>
- Ellegaard, E., & Wallin, J. A. (2015). The bibliometric analysis of scholarly production: How great is the impact? *Scientometrics*, 105, 1809-1831. <https://dx.doi.org/10.1007/s11192-015-1645-z>
- EPA. (2021). *Superfund, United States Environmental Protection Agency*. Retrieved from <https://www.epa.gov/superfund> & <https://www.epa.gov/vaporintrusion>
- Eslami, H., Ebadi, A., & Schiffauerova, A. (2013). Effect of collaboration network structure on knowledge creation and technological performance: the case of biotechnology in Canada. *Scientometrics*, 97(1), 99-119. <https://doi.org/10.1007/s11192-013-1069-6>
- Eurostat. (2018). Key figures on Europe – Statistics illustrated, 2018 edition. <http://dx.doi.org/10.2785/594777>
- Farhan, M. T., Darwiyanto, E., & Asror, I. (2019). Analysis of hubs and authorities centrality using probabilistic affinity index (PAI) on directed-weighted graph in social network analysis. *Journal of Physics*. <https://dx.doi.org/10.1088/1742-6596/1192/1/012005>
- Ferguson, C. C. (1999). Assessing risks for contaminated sites: policy and practice in 16 European countries. *Land Contamination and Reclamation*, 7(2), 87-108. Retrieved from <http://epppublications.com/Documents/07-2-1.pdf>
- Fillbrunn, A., Dietz, C., Pfeuffer, J., Rahn, R., Landrum, G. A., & Berthold, M. A. (2017). KNIME for reproducible

- cross-domain analysis of life science data. *Journal of Biotechnology*, 261, 149-156. <https://dx.doi.org/10.1016/j.jbiotec.2017.07.028>
- Fonseca e Fonseca, B., Sampaio, R. B., Fonseca, M. V. A. F., & Zicker, F. (2016). Co-authorship network analysis in health research: method and potential use. *Health Research Policy and Systems*, 14(34). <https://dx.doi.org/10.1186/s12961-016-0104-5>
- Freeman, L. C. (1978). Centrality in social networks: conceptual clarification. *Social Networks*, 1(3) 215-239. [https://doi.org/10.1016/0378-8733\(78\)90021-7](https://doi.org/10.1016/0378-8733(78)90021-7)
- Gadat, S., Gavra, I., & Risser, L. (2016). How to calculate the barycenter of a weighted graph. *Toulouse School of Economics*, working paper TSE-652. Retrieved from <https://www.tse-fr.eu/publications/how-calculate-barycenter-weighted-graph>
- Glanzel, W., & Schubert, A. (2005). Analyzing scientific networks through co-authorship. In: Moed H. F., Glanzel W., & Schmoch U. (Eds.), *Handbook of Quantitative Science and Technology Research* (pp. 257-276). Springer. https://doi.org/10.1007/1-4020-2755-9_12
- Gong, M., Li, G., Wang, Z., Ma, L., & Tian, D. (2016). An efficient shortest path approach for social networks based on community structure. *CAAI Transactions on Intelligence Technology*, 1(1), 114-123. <https://dx.doi.org/10.1016/j.trit.2016.03.011>
- Haythornthwaite, C. (1996). Social network analysis: An approach and technique for the study of information exchange. *Library & Information Science Research*, 18(4), 323-342. [https://doi.org/10.1016/S0740-8188\(96\)90003-1](https://doi.org/10.1016/S0740-8188(96)90003-1)
- JRC. (2014). Progress in the management of Contaminated Sites in Europe. <https://doi.org/10.2788/4658>
- Kamada, T., & Kawai, S. (1989). An algorithm for drawing general undirected graphs. *Information Processing Letters*, 31(1), 7-15. [https://doi.org/10.1016/0020-0190\(89\)90102-6](https://doi.org/10.1016/0020-0190(89)90102-6)
- Katz, J. S., & Martin, B. R. (1997). What is research collaboration? *Research Policy*, 26(1), 1-18. Retrieved from <https://EconPapers.repec.org/RePEc:eee:respol:v:26:y:1997:i:1:p:1-18>
- Kleinberg, J. M. (1999). Authoritative sources in a hyperlinked environment, *Journal of the ACM (JACM)*, 46(5), 604-632. <https://doi.org/10.1145/324133.324140>
- Knime. (2021). *Description of the "Network Analyzer" node*. Retrieved from <https://hub.knime.com/knime/extensions/org.knime.features.network/latest/org.knime.network.node.analysis.network.NetworkAnalyzerNodeFactory>
- Kolleck, N. (2013). Social network analysis in innovation research: using a mixed methods approach to analyze social innovations. *European Journal of Futures Research*, 1(25). <https://doi.org/10.1007/s40309-013-0025-2>
- Lander, B. (2013). Sectoral collaboration in biomedical research and development. *Scientometrics*, 94(1), 343-357. <https://doi.org/10.1007/s11192-012-0776-8>
- Laudel, G. (2002). Collaboration and reward. What do we measure by co-authorships? *Research Evaluation*, 11(1), 3-15. <https://doi.org/10.3152/147154402781776961>
- León, C. (2013). Authority centrality and hub centrality as metrics of systemic importance of financial market infrastructures. *SSRV platform*. <https://dx.doi.org/10.2139/ssrn.2290271>
- Marra, M., Emrouznejad, A., Ho, W., & Edwards, J. (2015). The value of indirect ties in citation networks: SNA analysis with OWA operator weights. *Information Sciences*, 314, 135-151. <https://doi.org/10.1016/j.ins.2015.02.017>
- McAlary, T., Provoost, J., & Dawson, H. (2011). Chapter 10 - Vapour intrusion. In Swartjes, F. (Ed.), *Dealing with Contaminated Sites – From Theory towards Practical Application* (pp. 409-453). Springer Publishers, Netherlands. Retrieved from <https://www.springer.com/gp/book/9789048197569>
- Melin, G., & Persson, O. (1996). Studying research collaboration using co-authorships. *Scientometrics*, 36(3), 363-377. <https://doi.org/10.1007/BF02129600>
- Merigó, J. M., & Yang, J. B. (2017). A bibliometric analysis of operations research and management science. *Omega*, 73, 37-48. <https://dx.doi.org/10.1016/j.omega.2016.12.004>
- Morel, C. M., Serruya, S. J., Penna, G. O., & Guimarães, R. (2009). Co-authorship network analysis: a powerful tool for strategic planning of research, development and capacity building programs on neglected diseases.

- PLoS Neglected Tropical Diseases*, 3(8), e501. <https://doi.org/10.1371/journal.pntd.0000501>
- National Research Council. (1997). *Innovations in Ground Water and Soil Cleanup: From Concept to Commercialization*. The National Academies Press, Washington DC. <https://doi.org/10.17226/5781>
- Newman, M. E. J. (2004). Co-authorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences of the United States of America*, 101(S1), 5200-5205. <https://doi.org/10.1073/pnas.0307545100>
- NICOLE. (2002). *Need for sustainable land management: Role of a risk assessment based approach*. Discussion paper, issue 2. Retrieved from <https://www.nicole.org>
- NICOLE. (2012). *How to implement sustainable remediation in a contaminated land management project?* NICOLE Sustainable Remediation Work Group report. Retrieved from <https://www.nicole.org>
- O'Madadhain, J., Fisher, D., Smyth, P., White, S., & Boey, Y. B. (2005). Analysis and visualization of network data using JUNG. *Journal of Statistical Software*, 1-25.
- Phillips, A. S., Hung, Y. T., & Bosela, P. A. (2007). Love Canal Tragedy. *Journal of Performance of Constructed Facilities*, 21(4). [https://doi.org/10.1061/\(ASCE\)0887-3828\(2007\)21:4\(313\)](https://doi.org/10.1061/(ASCE)0887-3828(2007)21:4(313))
- Ramadan, E., Alinsaif, S., & Hassan, R. (2016). Network topology measures for identifying disease-gene association in breast cancer. *BMC Bioinformatics*, 17(supplement 7), 474-543. <https://dx.doi.org/10.1186/s12859-016-1095-5>
- Rosas, R. S., Kagan, J. M., Schouten, J. T., Slack, P. A., & Trochim, W. M. K. (2011). Evaluating research and impact: A bibliometric analysis of research by the NIH/NIAID HIV/AIDS clinical trials networks. *PLOS One*, 6(3), e17428. <https://doi.org/10.1371/journal.pone.0017428>
- Sabidussi, G. (1966). The centrality index of a graph. *Psychometrika*, 31(4), 581-603. Retrieved from <http://hdl.handle.net/10.1007/BF02289527>
- Sonnenwald, D. (2007). Scientific collaboration. *Annual review of information science and technology*, 41(1), 643-681. <http://dx.doi.org/10.1002/aris.2007.1440410121>
- The Royal Society. (2011). Knowledge networks and nations: Global scientific collaboration in the 21st century. *RS Policy document 03/11*. Retrieved from https://royalsociety.org/-/media/Royal_Society_Content/policy/publications/2011/4294976134.pdf
- Tiwari, A., & Sekhar, A. K. T. (2007). Workflow based framework for life science informatics. *Computational Biology and Chemistry*, 31(5-6), 305-319. <https://dx.doi.org/10.1016/j.compbiolchem.2007.08.009>
- Tsay, M., & Shu, Z. (2011). Journal bibliometric analysis: a case study on the Journal of Documentation. *Journal of Documentation*, 67(5), 806-822. <https://doi.org/10.1108/00220411111164682>
- Valente, T. W. (2010). *Social networks and health: models, methods, and applications*. Oxford University Press, New York. <https://doi.org/10.1093/aje/kwq243>
- Viswanath, M. (2009). *Ontology-based automatic text summarization*. Master thesis, University of Georgia. Retrieved from https://getd.libs.uga.edu/pdfs/viswanath_meghana_200912_ms.pdf
- Wang, D. J., Shi, X., Mcfarland, D. A., & Leskovec, J. (2012). Measurement error in network data: a reclassification. *Social Networks*, 34(4), 396-409. <https://doi.org/10.1016/j.socnet.2012.01.003>
- Wasserman, S., & Faust, K. (1994). *Social network analysis: methods and applications* (Vol. 8). Cambridge University Press. <https://doi.org/10.1017/CBO9780511815478>
- Wood, J. A., & Porter, M. L. (1987). Hazardous Pollutants in Class II Landfills. *Journal of the Air & Waste Management Association - JAPCA*, 37(5), 609-615. <https://doi.org/10.1080/08940630.1987.1046625>
- Xutao, L., Michael, K., & Yunming, Y. (2012). HAR: Hub, Authority and Relevance Scores in Multi-Relational Data for Query Search. *Proceedings of the 2012 SIAM International Conference on Data Mining*. Retrieved from <https://epubs.siam.org/doi/abs/10.1137/1.9781611972825.13?mobileUi=0>
- Yang, C. H., & Heo, J. (2014). Network analysis to evaluate cross-disciplinary research collaborations: The Human Sensing Research Center, Korea. *Science and Public Policy*, 41(6), 734-749. <https://doi.org/10.1093/scipol/scu007>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).