Visual research landscape of surface water and groundwater interactions (1980 - 2017)

Dasapta Erwin Irawan[1], Budi Brahmantyo[1], Deny Juanda Puradimaja[1], Imam Priyono[1], and Ahmad Darul[2]

[1]*Faculty of Earth Sciences and Technology, Institut Teknologi Bandung, dasaptaerwin@outlook.co.id*
[2]*Faculty of Mining Engineering, Institut Teknologi dan Sains Bandung*

Abstract

Surface water is closely linked to unconfined groundwater, especially in Indonesia, where shallow aquifers are mostly composed of unconsolidated layers. To put Indonesia's research on surface water and groundwater interactions in to perspective, we did a scientometric study based on literature search using VosViewer and Scopus database. We analyzed a corpus of 281 papers to extract the hidden pattern of information using text mining algorithm. There are five clusters of subtopic that we gained from the dataset: modeling/simulation, contamination, water balance, and climate-land use. We also find multiple gaps in the literature that mostly related to local geological settings. Based on that, we endorse Indonesia authors to start looking at the unique socio-economic and biogeophysical features, and report it in form of scientific papers/reports and disseminate their key points online to get more international attention. Based on the results, we also push Indonesian scientific community to lead more international collaboration to promote local distinctive features using widespread collaboration tools online. Given the lack of funding, we also promote the usage of free and open source tools for any kinds of research initiatives.

Keywords: surface water, groundwater, interaction, gap analysis

1   Introduction

Indonesia, with its tropical setting, high rainfall, high humidity, and shallow unconsolidated aquifers, have made a close link between surface water (river water, lake water) with the surrounding groundwater (Irawan et al., 2016; Devi et al., 2017; Rahmadia et al., 2017; Rosada et al., 2017; Widodo, 2013, Satrio et al., 2012; Azy and Hadian, 2016; Putranto and Rde, and Taufik et al., 2017). We are interested to know how far people had been working on this bit in the world. We'll try to funneled the result to Indonesia's research in such subject.

Here we try to visualize the landscape of hydrology subject under the terms surface water and groundwater interactions. We will make a preliminary "gap

analysis" in the literature (Jennings 2000).

Our first idea was to use Scopus (S) data as the source of English-based papers and Google Scholar (GS) data as the source of Indonesian language-based papers. However, along the way we failed to harvest GS data due software crash.

2   Materials and methods

We urge the usage of open source applications for our work and we looked at VosViewer (Van Eck and Waltman, 2010) and CiteSpace (Chen et al., 2012) applications which have been continuously developed and have quite strong community. We also looked at Publish or Perish (PoP) applications (Harzing and Van der Wal, 2008) to harvest Google Scholar data. However the Citespace and PoP didn't work at the time of analysis. We suspected the crashes were caused by a conflict between Java and PlayonMac (an PC emulator for Mac computers) version with our Mac system.

The keywords that we used on Scopus was surface water AND groundwater AND interactions using title filter. We did the survey on 30 December 2017. The result may be slightly different as many new items would be indexed by Scopus in two weeks time, when we wrote the manuscript. We saved the literature database to RIS format before importing it to VosViewer. Dataset is available in bib, csv, and RIS format as supplementary materials.

VosViewer is a free tools to make visualization based on a set of corpus. In VosViewer we create network visualizations based on keywords and authors. Basically VosViewer reads the corpus (text) from the RIS file, count and sort them out, makes connections, and builds clusters from them.

We used the following procedures in VosViewer:

- **Procedure for keyword network**:

  - **Procedures 1 (unsupervised)**: Open VosViewer, Create a map based on bibliographic data (next) Load the RIS file (next) Type of analysis: Co-occurrence (next) Counting method: Full counting (next) Minimum number of occurrences of keyword: 5 (next) Number of keywords to be selected: 222 (next) Evaluate keywords (finish). We gained 222 keywords in the clusters.

  - **Procedures 2 (supervised)**: Open VosViewer, Create a map based on bibliographic data (next) Load the RIS file (next) Type of analysis: Co-occurrence (next) Counting method: Full counting (next) Minimum number of occurrences of keyword: 10 (next) Number of keywords to be selected: 90 (next) Evaluate keywords (finish). We gained 90 keywords in the clusters.

  - We chose procedures 2 with more sensible keywords clusters.

- **Procedure for authorship network**:

  - **Procedure 1 (unsupervised)**: Open VosViewer, Create a map based on bibliographic data (next) Load the RIS file (next) Type of analysis: Co-authorship (next) Counting method: Full counting (next) Minimum number of document of an author: 5 (next) Number

of keywords to be selected: 7 (next) Evaluate results (next) Show connected results only: No (finish). We gained seven authors with minimum of five documents in the clusters. We need to get more authors to see the pattern. Then we use procedure 2.

- **Procedure 2 (supervised)**: Open VosViewer, Create a map based on bibliographic data (next) Load the RIS file (next) Type of analysis: Co-authorship (next) Counting method: Full counting (next) Minimum number of document of an author: 3 (next) Number of keywords to be selected: 7 (next) Evaluate results (next) Show connected results only: No (finish). We gained 37 authors with minimum of three documents in the clusters.

3 Results and discussions

3.1 Keyword network

Initially we have 222 keywords in four clusters in the first trial (Figure 1 and Figure 2). Then we narrow downscale the keywords to 90 keywords in four clusters (Figure 3 and Figure 4) (see keyword cluster table as supplementary dataset). In the second trial charts, we see a closely linked clusters of text, that we dont see on the author collaborations map (we talk about the authorship side later one). There are five clusters (in clock-wise direction) which are connected at the center to the surface water and groundwater interactions:

1. Green cluster (GC) - subtopic: numerical/mathematical modeling, hydraulic properties, groundwater flow

2. Blue cluster (BC) - subtopic contamination: water quality, contamination, and contaminant transport

3. Red cluster (RC) - subtopic chemistry-isotope: water quality, water chemistry, and isotopes

4. Yellow cluster (YC) - subtopic water balance/budget: water balance/budget, recharge, irrigation

5. Purple cluster (PC) - subtopic climate-usage: water management, spatio-temporal analysis, vegetation

YC and PC are presumably belong to the same cluster, water balance or water budget. The keyword "water resource management" although it's included in YC, but the position is still secluded from the rest of the keywords. Does it mean that the papers in this category haven't been clearly linked (have a citation relationship) with papers in the other categories? The word "tropic" or "tropical" is not visible in the clusters. Does it mean we have less paper on tropical climate or tropical country or tropical hydrology/hydrogeology? We still have to closely understand the data structure.

More on it, we also don't see geophysical-related keywords. Does it imply that geophysical methods have not been closely linked to water resource mapping? We also miss "land subsidence" and "landslide", which has been the major

geohazard in many big cities. We only see "saline intrusion" in BC (Abidin et al., 2001; Marfai, 2007; Putranto and Kusuma, 2009).

In terms of geological system, we only see "wetlands" in the figure, but we don't see "peat". More systems like volcanic system, coastal system, or alluvium system are just an example of what we missed in figure.

Figure 1: Clustering based on keywords - first trial

Figure 2: Clustering based on keywords - first trial

3.2 Authorship network

Unlike the keyword network, we don't see a good links in between authors. We notice at least four clusters of dominant authors from: Australia, UK, US,

Figure 3: Clustering based on keywords - second trial
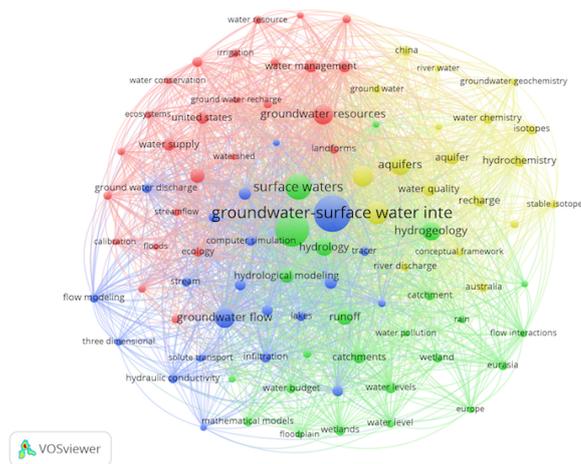
Figure 4: Clustering based on keywords - second trial

Taiwan and China. We need to dig up more on this since to uncover scientific patterns. Looking at the figure, we argue that this is another gap yang can be filled by Indonesian scientist. We understand that this paper is based on English language papers, but we are convinced that Indonesian authors can contribute more on this subject using Indonesia language. We recommend to future Indonesian authors to promote their local socio-economic-bio-geophysics originality.

The hydrological system of Indonesia is mostly composed of unconsolidated deposits forming unconfined aquifer systems. With such high rainfall at almost any part of region, we may expect many permanent and temporal rivers flowing on the valleys and lowlands. The changes of land-use across time is an-

5

other component that creates a unique interactions between surface water and groundwater. Those three components would build a complex system of water flowing and water transfer across the country.

We also want to promote collaboration between Indonesian authors at international level based on those originalities. There are many tools and communities now to start the effort with and lead a collaborative team.
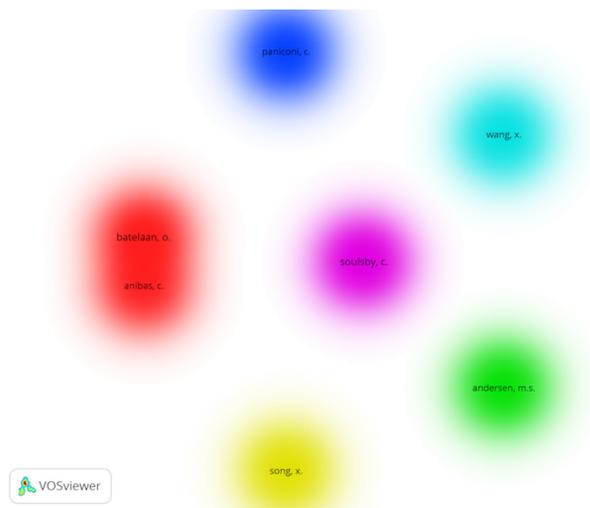

Figure 5: Clustering based on authors - first trial

Figure 6: Networking based on authors - first trial
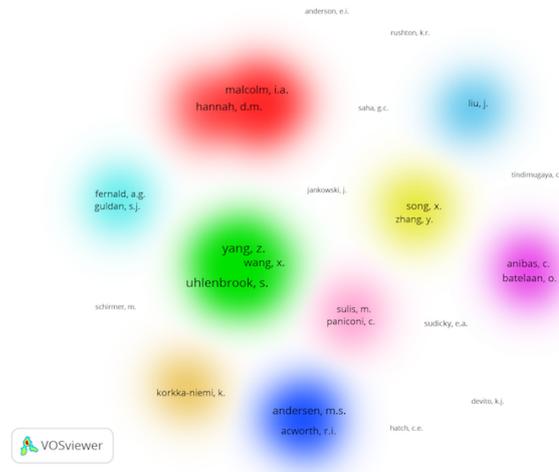
Figure 7: Clustering based on authors - second trial
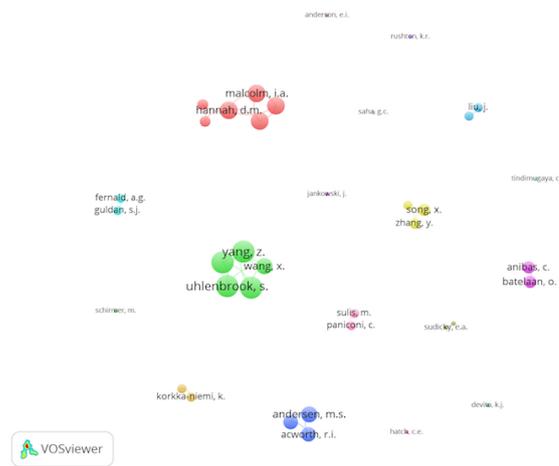
Figure 8: Networking based on authors - second trial

4   Concluding Remarks

Given the search strategy, we understand the bias that may arises from this paper, since we only accommodate English-based paper. Our next effort would be using another inclusive database to search Indonesian-based literatures. However, the results of this paper may be used as the basis to analyze the current status of papers authored by Indonesian scientists. Research institutions and universities, especially the ones that don't have access to commercial database like Scopus or Web of Science, may take advantages from this paper and the metadata that we share in the background.

Since we used free and open source tools to finish this paper, another re-

searchers could also replicate our method for their own purposes. It also may be used for non-engineering based fields.

Here are our annotation of important points:

- Many authors have contributed their knowledge to this subject and rapid growth started in early 2000's.

- Closely connected keywords but less collaboration (cross-country collaboration) among scientist in this field.

- There are weak (or underrepresented) papers to connect "water resource management" with other keywords.

- There are several preliminary-identified gaps in this corpus: geophysical method, tropical region, more water-related geohazards (like land subsidence and landslide), more systemic study based on local geological setting, such as: volcanic hydrogeological system, groundwater system in coastal area).

- Each gap is an opportunity for Indonesian authors to explore the originality of Indonesia that have been deeply covered by our own ignorance.

Acknowledgements

References

[1] Irawan, Dasapta, Adhi Priyambodho, Cut Rachmi, Dimas Wibowo, and Andita Fahmi. "Bibliometric study to assist research topic selection: a case from research design on Jakartas groundwater (part 1)." *Research Ideas and Outcomes Journal,* 2(2016): e9841, doi: 10.3897/rio.2.e9841.

[2] Devi, Meli Triana; Sulistyawati, Endah; Irawan, Dasapta Erwin; Rosada, Keukeu K., Vegetation of riparian areas along Cikapundung River, *Figshare Repository* (2017). doi: 10.6084/m9.figshare.5425714.v1.

[3] Rahmadia, Tantra; Sulistyawati, Endah; Irawan, Dasapta Erwin; Rosada, Keukeu K., Cikapundung watershed land cover characterization using SPOT-6 imagery, *Figshare Repository* (2017), doi:10.6084/m9.figshare.5425729.v1.

[4] Rosada, K., Sulistyawati, E., Irawan, DE., Ecoregion Cikapundung Watershed Based on The Presence of Escherichia coli – Poster Ecodevelopment 2017, *Figshare Repository* (2017). doi: 10.6084/m9.figshare.5479735.v1.

[5] Widodo, Lilik Eko. "Estimation of Natural Recharge and Groundwater Build up in the Bandung Groundwater Basin Contributed from Rain Water Infiltration and Inter-aquifer Transfer." *Procedia Earth and Planetary Science,* 6(2013), 187-194, url: available from ScienceDirect.

[6] Satrio, Satrio, S. Paston, Leong Chung Sum, and S. Syafalni. "Groundwater Dynamic and Its Interrelationship with River Water of Bandung Basin Using Environmental Isotopes (18O, 2H, 14C)." *Modern Applied Science,* 6(2012), no. 11, 49, url: available from CCSENET.

[7] Azy, Fikri Noor, and Mohamad Sapari Dwi Hadian. "Groundwater Characterization of Cihaur Watershed Basin, Batujajar and Adjacent, West Bandung District, West Java, Indonesia." *IOP Conference Series: Earth and Environmental Science,* vol. 29(2016), no. 1, 012027, doi: 10.1088/1755-1315/29/1/012027/pdf.

[8] Putranto, T. T., and T. R. Rde. "Hydrogeology of Semarang Demak Groundwater Basin: An overview and its challenges in preliminary groundwater flow modeling." *Proceedings of The 36th HAGI and 40th IAGI Annual Convention and Exhibition Makassar,* (2011), 26-29, url: available from ResearchGate.

[9] Taufiq, A., Hosono, T., Ide, K., Kagabu, M., Iskandar, I., Effendi, AJ., Hutasoit, LM., and Shimada, J, Impact of excessive groundwater pumping on rejuvenation processes in the Bandung basin (Indonesia) as determined by hydrogeochemistry and modeling *Hydrogeology Journal* (2017), doi:10.1007/s10040-017-1696-8.

[10] Jennings, Michael D. "Gap analysis: concepts, methods, and recent results." *Landscape ecology 15,* no. 1(2000), 5-20, url: available from Semantic Scholar.

[11] Van Eck, N. J., and Waltman, L., Software survey: VOSviewer, a computer program for bibliometric mapping, *Scientometrics,* 84(2), 523-538, doi: 10.1007/s11192-009-0146-3.

[11] Chen, C., Hu, Z., Liu, S., and Tseng, H., Emerging trends in regenerative medicine: a scientometric analysis in CiteSpace, *Expert opinion on biological therapy,* 12(5), 593-608, doi: 10.1517/14712598.2012.674507.

[12] Harzing, Anne-Wil K., and Ron Van der Wal. "Google Scholar as a new source for citation analysis." *Ethics in science and environmental politics 8,* no. 1(2008), 61-73, url: available from Inter Research Science Center.

[13] Abidin, Hasanuddin Z., Rochman Djaja, Dudy Darmawan, Samsul Hadi, Arifin Akbar, H. Rajiyowiryono, Y. Sudibyo et al. "Land subsidence of Jakarta (Indonesia) and its geodetic monitoring system." Natural Hazards 23, no. 2 (2001): 365-387, doi: href https://doi.org/10.1023/A:10111446010.1023/A:101114460.

[14] Marfai, M.A. and King, L., Monitoring land subsidence in Semarang, Indonesia Environmental Geolology 53(2007), 651, doi: 10.1007/s00254-007-0680-3.

[15] Putranto, TT. and Kusuma, KI., Permasalahan Airtanah Pada Daerah Urban, Teknik 30(2009), 1, 48-56, doi: 10.14710/teknik.v30i1.1824 .