What do we Know about the Grand Ethiopian Renaissance Dam and the Abay Basin: A Review Using Text Mining and Natural Language Processing Techniques

Wuletawu Abera[1], Lulseged Tamene[1], Melkamu Beyene[2], Bryce Gallant[3], Nigussie Haregeweyn[4], Yihun Dile[5], Ayele A. Fenta[6], Hailemariam Meaza[7], Kifle Woldearegay[8], Biadgilgn Demissie[7]

Abstract

The Grand Ethiopian Renaissance Dam (GERD) is a mega-hydropower plant under construction on the Upper Blue Nile (UBN) River (Abay) in Ethiopia. The project attract attention from both the political and scientific communities thus many studies are conducted considering its different aspects in the last nine years. Studies are multidisciplinary ranging from hydrology, geology, and ecosystem, and socio-economics, politics, legal, etc. of the GERD and the basin. Until July 2020, more than 250 peer-reviewed papers and more than 150 gray literature (thesis and reports) are published. To distil and understand the knowledge and information created in these papers, researchers often need to conduct systematic review of studies. However, the traditional approach of doing literature review is not efficient and can be complicated as the volume of literature to scan, read and review is large. In this study, we aim to stock take and assess the knowledge and information generated by scientific community since the start of the GERD construction using text mining (TM) and natural language processing (NLP) algorithms. The TM results explore and appraise the key knowledge, findings, and enable to discern topics of key discussions regarding GERD and its basin in general. The study focused on: 1) conducting basic bibliometric statistics and trends in the frequency of key topics over time related to the GERD; 2) use machine learning particularly TM and NLP approaches for scientific review from structured and unstructured documents and transform the traditional review process; 3) employ the network of keywords to identify which thematic area or topic is the most studied topic related to GERD and vice versa; 4) identify negative, positive, and neutral sentiments towards GERD project; 5) identify research gaps and guide future research directions regarding the GERD and the basin.

1 International center for Tropical Agriculture (CIAT), Addis Ababa, Ethiopia; Email: wuletawu.abera@cgiar.org

2 School of Information Science, Addis Ababa University, Ethiopia

3 International Water Management Institute (IWMI), Sri Lanka

4 International Platform for Dryland Research and Education, Tottori University, Tottori, Japan

5 College of Agriculture and Life Sciences, Texas A&M University, College Station, TX, USA

6 Arid Land Research Center, Tottori University, 1390 Hamasaka, Tottori, 680–0001, Japan

7 Mekelle University, Department of Geography and Environmental studies, Mekelle, Ethiopia

8 Mekelle University, Department of Earth Science, Mekelle, Ethiopia

2