The Use of Big Data in Water Resources Management

Sara Bouziane[1], Badraddine Aghoutane[1], Aniss Moumen[2], Ali Sahlaoui[3], Anas El Ouali[4]
[1]*Informatics and Applications Laboratory, Science Faculty of Meknes, Moulay Ismail University, Meknes, Morocco*
[2]*Ibn Tofail University, National School of Applied Sciences, Kenitra, Morocco*
[3]*Laboratory of Geo-Engineering and Environment, Faculty of Sciences, Moulay Ismail University, Meknes, Morocco*
[4]*Department of Environment, Functional ecology and environmental engineering laboratory, Faculty of Sciences and Technology, Sidi Mohamed Ben Abdellah University, Fez, Morocco*

Keywords: Big Data Analytics, Water Resources Management, Information System, IoT.

Abstract: Water management has become an essential vector in the Moroccan government policy since independence. The volume of the water data, and the diversity of the actors, lead us to think about new methods of analysis to manage water resources with efficiency. The development of new technologies, such as Big Data, is an essential tool for assuring this management. Researchers have used these technologies to develop several architectures and algorithms to deal with the water scarcity concern. In this paper, we identify research works related to the topics of water and Big Data, and we discuss the different proposed architectures, according to a review of systemic exploratory literature. In the end, we draw up our perspectives.

1 INTRODUCTION

We found in recent years much talk in public debate and the social sciences about Big Data and IoT technologies (Elhassan et al., 2020). That provides methods to collect, manage, and process a large volume of data that ensures the management of large and diversified information systems such as water data.

In Morocco, water has been considered an essential component of the economy since independence (Hafed et al. 2018). In this perspective, the Water Basin Agency (ABH), as an independent public organization, is implementing a water management strategy with the various actors of the water sector at the level of the watersheds following the water strategy coordinated and monitored by the Directorate General of Water (Moumen et al. 2016).

And all over the world, due to increasing water demand, climate, and hydrological deficit, water resource managers have started to look for practical strategies for water resource management.

2 LITERATURE REVIEW

2.1 Methodology

The literature review refers to a process and reporting structure to classify and identify research and results published to date for a given topic (in our case Big Data in the field of water resources) (Moumen et al. 2016).

To conduct our Literature Review, as explained in Figure 1, we have identified 217 references in the various databases and have stored them in the Zotero library.

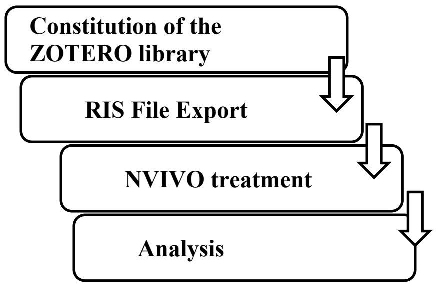

Figure 1: Working methodology.

Bouziane, S., Aghoutane, B., Moumen, A., Sahlaoui, A. and El Ouali, A.
The Use of Big Data in Water Resources Management.
DOI: 10.5220/0010728000003101
In *Proceedings of the 2nd International Conference on Big Data, Modelling and Machine Learning (BML 2021)*, pages 38-44
ISBN: 978-989-758-559-3

We have created this number of references using the method shown in Figure 2.

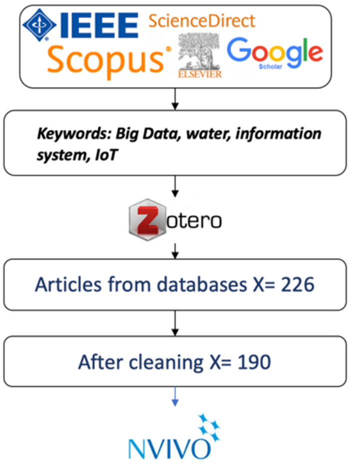

Figure 2: Literature review process.

2.2 Meta Analyzes

2.2.1 Analyze of References

Table 1 below represents the number of references by bibliographical category.

Table 1: References categories.

Type of references	Number
Journal articles	110 (57,9%)
Conference papers	76 (40%)
Chapters & books	4 (2,1%)
Total	190

We have collected these references from various databases, as shown in Figure 3 below. Scopus occupies first place with 79 articles, followed by the Institute of Electrical and Electronics Engineers (IEEE) with 64 references, and ScienceDirect with 40 documents. We have collected the other documents (7 references) from the Google Scholar database.

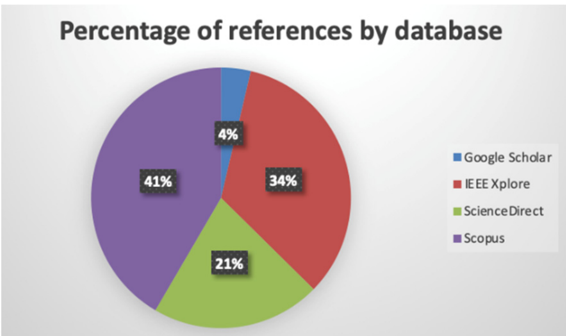

Figure 3: Percentage of references by databases.

Table 2 below presents the complete list of journals containing articles. We can notice that most articles are published in the journals "Advances in Intelligent Systems and Computing" and "Environmental Modelling & Software".

Table 2: Research papers per journals.

Journal	Nb. Articles
Advances in Intelligent Systems and Computing	6
Environmental Modelling & Software	5
Procedia Computer Science	4
Procedia Engineering	4
Water (Switzerland)	4
Communications in Computer and Information Science	3
Journal of Cleaner Production	3
Arabian Journal of Geosciences	2
Computers and Electronics in Agriculture	2
IEEE Access	2
Journal of Petroleum Science and Engineering	2
Sustainable Computing: Informatics and Systems	2
Unassigned	2
Water Research	2
Advances in Water Resources	1
Agricultural Systems	1
Agricultural Water Management	1
Applied Geography	1
Applied Sciences (Switzerland)	1
Array	1
Biosystems Engineering	1
Computational Geosciences	1
Computer Communications	1
Computer Networks	1
Computers and Geosciences	1
Computers, Materials and Continua	1
Concurrency Computation	1
Current Opinion in Biotechnology	1
Dili Xuebao/Acta Geographica Sinica	1
E3S Web of Conferences	1
Environmental science & technology	1
Environmental Science and Technology	1
Environmental Technology & Innovation	1
Environmental Technology and Innovation	1
Field Crops Research	1
Future Generation Computer Systems	1
Groundwater	1
Handbook of Environmental Chemistry	1
Hydrology and Earth System Sciences	1
HydroResearch	1
IEEE Latin America Transactions	1
IEEE Sensors Journal	1
IEEE Transactions on Big Data	1
IEEE Transactions on Industrial Informatics	1
IEEE Transactions on Sustainable Computing	1
International Journal of Embedded Systems	1
Irrigation and Drainage	1
Jilin Daxue Xuebao (Diqiu Kexue Ban)/Journal of Jilin University (Earth Science Edition)	1
Journal of Ambient Intelligence and Smart Environments	1
Journal of Coastal Research	1
Journal of Environmental Management	1

Table 2: Research papers per journals (cont.).

Journal	Nb. Articles
Journal of Hohai University	1
Journal of Marine Systems	1
Journal of Sustainable Development of Energy, Water and Environment Systems	1
Journal of the Chinese Institute of Civil and Hydraulic Engineering	1
Lecture Notes in Electrical Engineering	1
Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST	1
Linye Kexue/Scientia Silvae Sinicae	1
Meitan Xuebao/Journal of the China Coal Society	1
Modeling and Optimization in Science and Technologies	1
Nongye Jixie Xuebao/Transactions of the Chinese Society for Agricultural Machinery	1
Open Agriculture	1
Remote Sensing	1
Remote Sensing of Environment	1
Research of Environmental Sciences	1
Science of the Total Environment	1
Sensors	1
Sensors (Switzerland)	1
Sensors and Actuators B: Chemical	1
Shuili Fadian Xuebao/Journal of Hydroelectric Engineering	1
Shuili Xuebao/Journal of Hydraulic Engineering	1
SN Applied Sciences	1
Sustainability (Switzerland)	1
Sustainability Science	1
Sustainable Cities and Society	1
Utilities Policy	1
Water Resources and Industry	1
Water Resources Management	1
Water Science and Technology	1
Water Security	1
Wiley Interdisciplinary Reviews: Water	1
Total	**110**

The following Figure 7 shows that most of the sources in our dataset were published in the years 2020 and 2021. Also, the year 2015 was the start point of publishing works relating to water and Big Data.

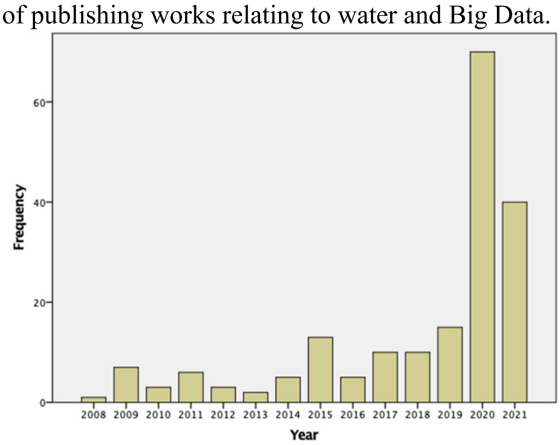

Figure 7: Number of sources per year.

Table 3 classifies the number of papers in our dataset by language, and we can see that 188 references are written in English, and only two sources are written in French.

Table 3: Research sources per language.

Language	Number
English	188
French	2
Total	190

2.2.2 Word Analysis

We present the most commonly used and treated words in our documents in a word cloud as a visual representation demonstrated in Figure 4.

At first glance, we note that 'Water', 'data', 'big', 'management', and 'system' are the most common words in our references. That refers to the central issue of this paper, which is the use of a Big Data system in water resources management.

Figure 4: Word cloud.

That is also illustrated in Figure 5 next, which represents the most frequent word occurrences. We notice that the predominantly used word are 'water', 'data', 'big', 'system', and 'management', which highlights the context of our paper.

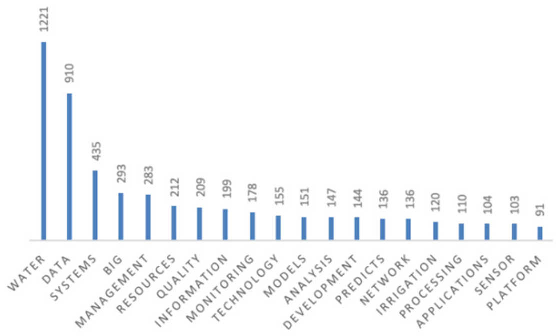

Figure 5: Most frequent words.

Figure 6 shows the least important commonly used words, which are ‘IoT’, ‘Lake’ (57), ‘Applying’, ‘Indices’ (56), ‘Groundwater’, ‘Objective’, ‘Conservancy’, ‘Demand’ (52), and finally ‘Challenges’ and ‘Implementation’ (51).

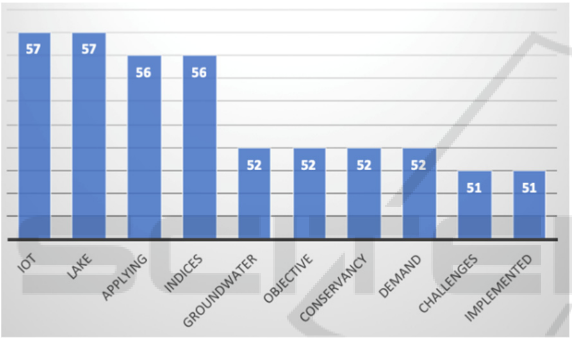

Figure 6: Least frequent word occurrences.

2.2.3 Grid Analysis

This thematic analysis provides a basis for data comparison to find out approaches in behaviour like domain, topic, technology, etc.

Our analysis allowed us to have 18 topics, as described in Table 4.

Table 4: List of relevant topics and associated terms.

Topic	**Words**
Big Data	"Big Data" processing analysis storage analytics "data management" visualization integration predicting
IoT	IoT sensor "internet of things" "remote sensing" real-time
Decision	decision "decision tool" "decision-makers" DDS SDSS "business intelligence" "decision making"
BDA tools	implementation spark apache Hadoop NoSQL Map Reduce framework
Analysis	Statistic exploratory analysis algorithm "machine learning"
Information system	"information system"
Cloud	"cloud computing"
Architecture	Architecture
Algorithm	Algorithm Algorithms
Computing	platform system computing
Water	water Groundwater hydrological wastewater
Climate	Climate
Agriculture	Agriculture
Urban	Urban
Energy	Energy
Food	Food

Analysis of the domains of application in Table 5 next shows that 180 documents deal with the topic of water, 47 documents treat the subject of decision-making, and the domains of agriculture (22), urban (26), energy (21), food (7), and climate (13).

Table 5: Domain of application.

Water	180
Decision	47
Agriculture	22
Urban	26
Energy	21
Food	7
Climate	13

As shown in Table 6, the commonly treated technologies are those associated with “Big Data” (178), followed by “Computing” (154), “Architecture” (25), “Algorithm” (32), “IoT” (92), “BDA Tools” (93), “Information System” (10), “Analysis” (100), and “Cloud computing” cited in 12 references.

Table 6: Technology and statistics.

Big Data	178
Computing	154
Architecture	25
Algorithm	32
IoT	92
BDA Tools	93
Information System	10
Analysis	100
Cloud	12

Table 7 below represents the crossover of the most cited technologies with fields of application. We notice that Big Data, Computing, IoT, and BDA Tools are the most generally cited technologies compared to others in the water area. That proves that we can move considerably in our research work in the Big Data application. We also notice that 32 references are talking about algorithms in the water field and 23 present architectures in this field of application.

Table 7: Grid of cited technologies and domains of application.

	Algorithm	Analysis	Architecture	BDA Tools	Big Data	Cloud Computing	Computing	Information system	IoT
Agriculture	3	7	3	9	19	1	17	1	11
Climate	1	4	2	5	11	1	12	0	6
Decision	10	27	7	19	45	6	39	4	25
Energy	5	12	6	10	21	1	19	1	9
Food	0	2	0	3	5	0	5	0	2
Urban	3	14	1	16	24	0	25	1	11
Water	32	97	23	87	169	10	145	10	88

3 DISCUSSION

At the end of this review, and based on the above crossover, we discuss and compare the different architectures applied in water management.

3.1 Theoretical Architectures

Using topics Architecture/Water, we found articles presenting Big Data architectures for water management. Following, we describe the three must-referenced results.

3.1.1 Big Data Open Platform for Water Resources Management

In this paper, the authors present the first part of a Big Data Open Platform that can manage and process a massive volume of water data collected from different sources and anticipate and avoid catastrophic situations such as floods and droughts. This part is developed and based on J2EE technology and provides better decision-making relating to the management of natural resources.

The presented model is composed of 9 blocks, as shown in Figure 7 below.

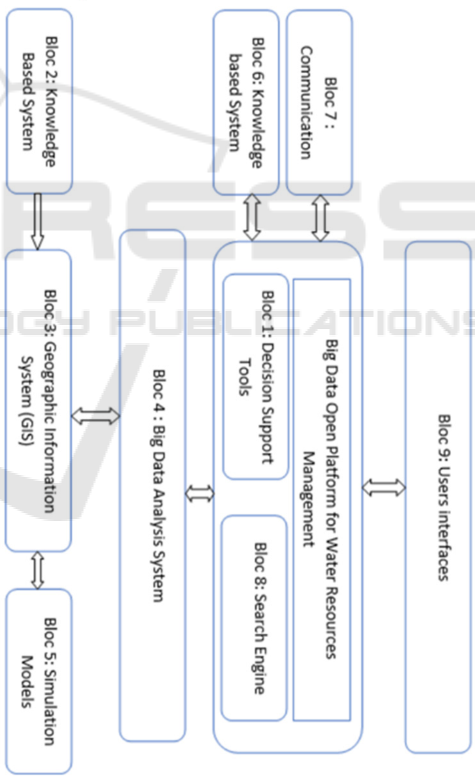

Figure 7: Big Data Open Platform for Water Resources Management proposed by Ridouane Chalh et al. (Chalh et al. 2015).

3.1.2 Application of Big Data in Water Ecological Environment Monitoring

In this paper, the authors present a Big Data system for China's water ecological data management, which the primary services are: the control of measurement of the water environmental data so we can use them in environmental protection projects; the government of the nutritional value of water bodies, so it helps in fisheries; and also it serves in the domains of the water conservancy engineering and coastal engineering by providing data support for environmental protection strategies of water conservancy projects and coastal projects.

Figure 8 next represents the presented model and shows the processing of the local data based on cloud computing and artificial intelligence that provides classification, count, and visualization of these data.

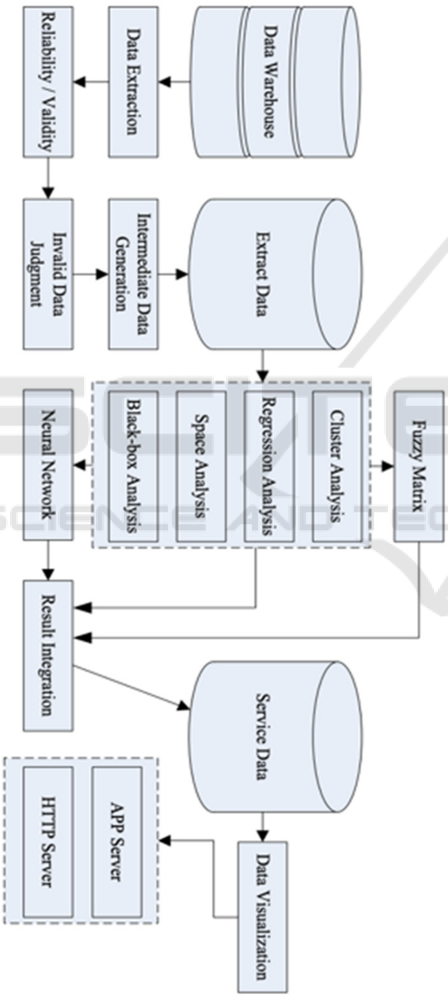

Figure 8: Flow Chart of Data Processing presented by Huanchun Ma et al. (Huanchun Ma et al. 2020).

The authors also presented some problems of water environmental data and the corresponding countermeasures.

3.1.3 A Pilot Infrastructure for Searching Rainfall Metadata and Generating Rainfall T Product using the Big Data of NEXRAD

In this paper, Bong-Chul Seo et al. (Moumen et al. 2016) present the pilot framework developed by The Iowa Flood Center (IFC) and known as IFC-Cloud-NEXRAD. The main objective of this infrastructure is to explore rainfall metadata and generate rainfall products over the Iowa domain. The framework is based on the NEXRAD radar Level II and data are accessible via cloud storage, which allows researchers to ready access to the radar data.

Figure 9 below presents the architecture of the IFC-Cloud-NEXRAD system.

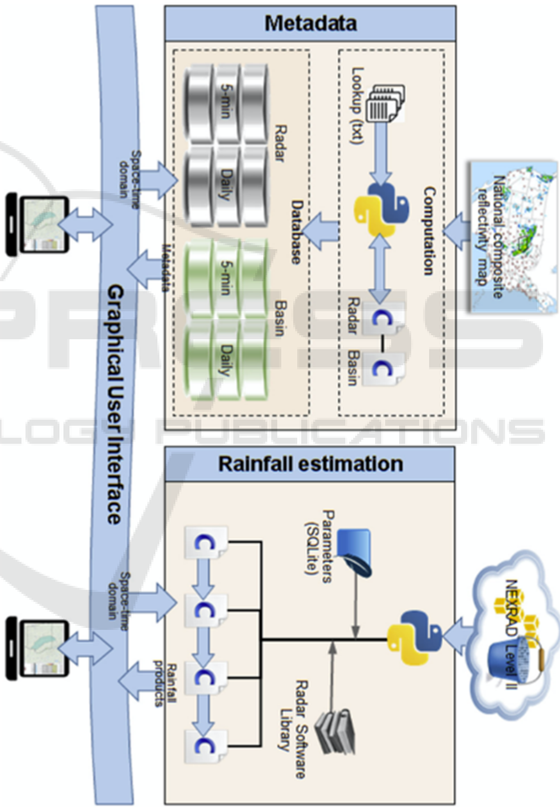

Figure 9: IFC-Cloud-NEXRAD framework architecture.

3.2 Comparison of Proposed Architectures

Table 8.

	Modules	Statut	Advantages	Limits
Architecture 1 Chalh et al. (2015)	▪ Users Interfaces ▪ Search Engine ▪ Decision Module ▪ Knowledge-Based System ▪ Big Data Analysis module ▪ Geographic Information System (GIS) ▪ Simulation Models ▪ Communication module ▪ Computing and processing module	The platform is in the design stage.	▪ Presence of a decision-making module ▪ Integration of the geographic module ▪ Presence of a search engine	▪ Architecture does not follow a process of data flow.
Architecture 2 Huanchun Ma et al. (2020)	▪ Big Data Analysis Module ▪ Geographic Information System (GIS) ▪ Data warehouse ▪ Visualization module ▪ Processing module ▪ User interface ▪ Decision-making ▪	The system is available.	▪ Neural network system and decision-making ▪ Data mining	▪ The automation of data acquisition still needs to be developed ▪ Data management is not secure enough ▪ The data mining scheme needs to be improved
Architecture 3 Bong-Chul Seo et al. (2019)	▪ Computing and processing module ▪ Users interface ▪ Space-time research ▪ Big Data Analysis module ▪ Decision and estimation module ▪ Geographic Information System (GIS) ▪ Radar ▪ Visualization Module	The framework is developed and could be transferrable to other geographic regions and	▪ Ready access to NEXRAD radar data ▪ Provides rainfall products ▪ Based on cloud computing ▪ Estimation of rainfall metadaa	▪ The maximum length of period for requesting products is limited to 30 days.

4 CONCLUSION

This work summarizes our exploratory study by identifying and discussing the different research works in the water domain and Big Data technologies.

In terms of perspectives, the next step is to study different proposed machine learning algorithms and models for collecting data relating to water and utilize these results to carry out the prototyping of a Big Data system for water management, according to users' expectations.

REFERENCES

Jamal Elhassan et al., 2020. *Big Data Analytic Architecture for Water Resources Management: A Systematic Review, In Proceedings of the 4th Edition of International Conference on Geo-IT and Water Resources 2020, Geo-IT and Water Resources 2020 (GEOIT4W-2020). Association for Computing Machinery, New York, NY, USA, Article 19, 1–5.*

Hafed, K. et al. 2018. *"Decisional Information Systems of the Public Actors in Moroccan Oasis Zones: Case Study Draa-Tafilalet Region", International Conference on Intelligent Systems and Computer Vision (ISCV 18), doi:10.1109/isacv.2018.8354047.*

Moumen, A. et al., 2016. *Web Services for Sharing and Reuse Water Data In Morocco. Proceedings of the Mediterranean Conference on Information & Communication Technologies 2015, 297–302. doi:10.1007/978-3-319-30298-0_31*

R.Chalh et al., June 2015. *Big Data Open Platform for Water Resources Management, International Conference on Cloud Technologies and Applications (CloudTech), Marrakech, Morocco.*

Huanchun Ma et al., 2020. *IOP Conf. Ser.: Mater. Sci. Eng. 750 012044.*

Bong-Chul Seo et al., 2019. *A Pilot Infrastructure for Searching Rainfall Metadata and Generating Rainfall T Product using the Big Data of NEXRAD, Environmental Modelling and Software 117 (2019) 69–75.*