ISSN (e): 2250 – 3005 || Volume, 08 || Issue, 8|| August – 2018 ||
International Journal of Computational Engineering Research (IJCER)

A Survey on Machine Learning Methods in GIS Based Water Modeling

J.Charles[1], G.Vinodhini[2], R.Nagarajan[3]
1Research scholar, Department of Computer and Information sciences,
2 Assistant Professor, Department of Computer science and Engineering,
3 Assistant Professor, Department of Computer and Information sciences,
Annamalai University, Annamalai Nagar.
Corresponding Author: J.Charles[1]

ABSTRACT
In the most recent few years, in addition to analyzing the time series data, space domain analytics has gained equal importance. Geographical Information Systems play a dominant part in analyzing and visualizing spatial data. The purpose of this paper is to provide an introduction to geographic information systems and a research background for information systems researchers on water modeling. In particular, this paper summarizes in the domain of GIS based water modeling. In addition, the applicability of machine learning in GIS based water modeling is discussed. Various tools used for GIS and machine learning are also provided. This survey shows that the existing standards, tools, GIS technologies and machine learning methods have been used effectively in the hydrological domain.
KEY WORDS : GIS; machine; learning; water, hydrology.

Date of Submission: 20-07-2018 Date of acceptance: 04-08-2018

I. INTRODUCTION

The association between water and the economy is so gripping that decisions about water are hardly ever delayed. Such decisions which are not conversant nearly all the time have unintentional consequences, with major impact on the environment and society. Changing climates are creating water shortages and changing flood patterns (Amir H.H et al., 2018; Chen.Q et al., 2016; Davoodi.M et al., 2015; Bodeaux.M et al., 2000). It is rare to find one level of government with sole jurisdiction over water management. Management is needed because water is not distributed evenly in space and time. Understanding its distribution can lead to solutions when water is temporarily unavailable.

Due to the changeable climatic conditions, universally in current years, the government has to seek out for apposite solution which can offer for the finest water management. In this view, geographic information systems will provide the goal. GIS is a computer system that provides accessibility, and allows simultaneous applicability of multiple issues (King T.V et al., 2018; Lee M.J et al., 2015; Lee M.J et al, 2004). GIS knowledge can superimpose different maps based on various aims, and also provide a comprehensive survey in water management. Therefore, the need of GIS tools for analyzing water management is essential and unavoidable. Till now, diverse types of models, including expert knowledge, statistics, analytical and machine learning have been used for water spatial modeling. Among these, the efficiency of different machine learning-based models has been approved in several studies (Naghibi, S. A et al., 2015; Saro.L et al., 2018). Machine learning can computerize, abridge and perk up many aspects of water modeling. This paper focuses on survey of various literatures that exists on proving efficiency of different machine learning based GIS models for water management.

The remainder of the paper is organized as follows, Section 2 describes about GIS in water modeling, and Section 3 is about Machine learning in GIS. Section 4 is on tools used for GIS and Machine learning. Section 5 describes Applications of GIS in water modeling. Section 6 concludes our survey.

II. GIS IN WATER MODELING

Geographic information system concepts and technologies are applied expansively in current water resources engineering, design, planning and operations and are changing the way these behavior are accomplished. GIS has become a progressively more significant means for understanding and dealing with the critical problems of water and related resources management in our world. GIS concepts and technologies help us collect and organize the data about such problems and understand their spatial relationships. GIS analysis capability offer methods for modeling information that contribute to sustaining decisions for resource management over a wide range of balance, from local to global. GIS provides a way for visualizing resource characteristics and thus improving the understanding in support of decision making. Many studies used integrated geographic information system methods for water management. (Heddam et al., 2016b; Heddam, S et al., 2016c; Clapcott et al., 2013; Baiyinbaoligao et al., 2011; Amiri B.J et al., 2009).

King T.V et al., 2018; Lee .S. et al., 2014 and Nikoo, M. R. et al., 2015 employed GIS functions to provide extensive means for developing surface water hydrologic models and operations. They said that the watershed runoff processes are inherently spatial in character so there is a strong motivation to use GIS tools to organize the data and formulate hydrologic models. Their research concludes that surface water hydrology is perhaps the area for which GIS has been applied effectively in the water resources and environmental field. They used DEMs, TINs, DLGs, digital soil and land use data, radar-rainfall and satellite imagery, real-time gage reporting, and the GIS software to process these has contributed to an increased awareness of the spatial distribution of hydrologic processes.

Shokoohi.M et al., 2017; Şener.S et al., 2017; Zhang et al., 2010; Wu Z.Y, et al., 2003 and Rossman L.A 2000 applied GIS in water supply distribution systems and components, aspects of their design and operation, and GIS concepts and tools which support these activities. These researchers also applied GIS tools for demand estimations, network design, and system operations. GIS spatial data management and analysis tools enable these functions.

Radin H.A , 2006; Rindahl.B 2013; and Maidment D.R, 2004 described that GIS has found extensive application for groundwater assessments as there are many types and large amounts of data involved. They also stated that proper evaluation of groundwater resources requires thorough hydrologic, geologic, and hydraulic investigations. Their work were based on the most popular computer model of the numerical type, the modular finite-difference groundwater flow model (MODFLOW) developed by the US Geological Survey. MODFLOW simulates groundwater flow in aquifer systems using the finite-difference method

III. MACHINE LEARNING IN GIS

The Machine learning technique is a swiftly emerging area of predictive modeling that is concerned with identifying structure in complex, often nonlinear, data and generating accurate predictive models. Machine learning approaches regularly demonstrate superior control for solving multifaceted relations. Machine learning approaches are not limited to the conventional assumptions normally used with traditional and parametric approaches (Sener.S 2017; Parsaie.A. 2018; Jaddi.N. S. 2017, Azamathulla H.M., 2016). Conventional modeling methods are normally based on statistical assumptions, data requirements and regularly employ linear modeling approaches that are inconsistent with usual processes that happen in the landscape (Davoodi M et al., 2015; Naghibi et al., 2015; Razandi et al., 2015; Clapcott et al., 2013; Manap et al.; 2012). Regression tree models are such examples of machine learning techniques that offer alternative approaches to the traditional methods of prediction.

Elith et al., 2008; Leathwick et al.; 2008 and De ’ath et al., 2010 proved boosted-tree algorithm as an influential approach for water modeling study. A study that used the boosted-tree algorithm for flood mapping was also conducted by Coltin et al., 2016. Lee et al., 2014, Lee M.J et al., 2015 and Lee 2015 showed the effective results of the analysis of the water modeling in the Seoul area using artificial neural network.

A number of studies have applied the random forest classifier for flood mapping for which other methods are integrated. Parsaie.A et al., 2018 applied machine learning models such as classification and regression tree techniques in GIS water modeling and groundwater studies (Chen , 2016; Elith et al., 2008). Multivariate statistical methods were used to study water quality assessment and spatiotemporal variations in water quality for Melen River water quality data by Koklu et al., 2010.

Stream water total nitrogen was predicted from land cover and human population density for 21 river basins using artificial neural networks and multiple linear regression by Amiri et al., 2009. Amiri et al., 2008 have also studied the linkage between land use and buffer zone with water quality using neural networks.

Amir.H et al., 2018 and Azamathulla, H.M et al., 2016 investigates the performance of artificial intelligence techniques including artificial neural network, group method of data handling and support vector machine (SVM) for predicting water quality components of Tireh River located in the southwest of Iran. They tested different types of transfer and kernel functions to develop the ANN and SVM. Reviewing the results of ANN

and SVM , they indicated that both models have suitable performance for predicting water quality components. Shokoohi et al., 2017 managed the water quality of a water supply system. They considered this an optimization problem and used modern optimization methods to solve it. Heddam et al., 2017 a utilized artificial neural networks for predicting the water quality components in several case studies. Emamgholizadeh et al., 2014 used multilayer perceptron, radial basis network and an adaptive neurofuzzy inference system for water quality components of Karoon River. They stated that all applied models have suitable performance for prediction of water quality components; however, the multilayer perceptron model was slightly more accurate. Zhang et al. 2010 introduced a new neural network approach for water allocation. They considered water quality as one of the main factors in their approach.

Nikoo et al., 2015 developed a probabilistic support vector machines model associated with GIS technique for planning the classification and distribution of surface and groundwater water in Iran. They stated that the use of these two methods would provide accurate information for feasibility studies of water conservation projects.

IV. TOOLS

GIS is extremely influential for addressing diverse water resources issues such as water quality, ground water movement, ground water contamination, river restoration, flood prediction and management, and etc. on a local, regional, national or even global scale (Parsaie.A. 2018; Sener.S 2017; Jaddi.N. S. 2017, Azamathulla H.M., 2016; Hammouri, N. et al., 2007). It could be used for different approaches such as analyzing the current situation, modeling and stimulating different scenarios for predicting the future, projecting new information, and enhancing decision making and watershed management (Chen et al., 2004). New developments in geographic information systems are providing a stronger framework for synthesizing geospatial and temporal water resources data into a single information system.

1.1. GIS tools

- Arc Hydro is a data model and toolset for water resources application, built within the ArcGIS software system. Arc Hydro includes a network of rivers and streams, a means for defining drainage areas and linking them to the network, connections to time series recorded at gaging sites, and a three-dimensional model for channel shape (Colti et al., 2016).
- Quantum GIS is an open-source and cross-platform GIS tool that allows the user to create, manipulate, analyze and visualize geographical data. Quantum GIS represents the maps in vector format or raster data (Ilayaraja, K, et al., 2015).
- gvSIG is another open source GIS tool that aids in solving sophisticated management problems by capturing, storing and handling the geographical data. It also can handle vector and raster data (Rossetto, R. U. D. Y. et al., 2013).
- PostGIS tool scores high on performance and is pretty fast when loading large volume of raster data. GRASS is the GIS tool well suited for scientific problems. It also has support for image processing and handling of graphical data. This tool is specialized for spatial modeling, maps and graphic production (Swain, N. R et al., 2015).

1.2. Machine learning tools

- Flink-ML is a machine learning library currently in development for the Flink platform. It supports implementations of Logistic Regression, k-Means Clustering, and Least Squares (Lee S et al., 2015).
- Weka 3 is a collection of machine learning algorithms for data mining tasks.
- H20 is an open source Environment for predictive analytics platform.
- Mahout is a framework for building scalable algorithms for shallow learning.
- Theano is a Python based numerical computation library for Deep and shallow learning.
- R is an Environment for statistical shallow learning.
- Torch 7 is a scientific computing framework with wide support for machine learning algorithms.

Applications

GIS find its application on diverse water resources issues such as water quality, ground water movement, ground water contamination, river restoration, flood prediction and management, and etc. on a local, regional, national or even global scale (Parsaie.A. 2018; Sener.S 2017; Jaddi.N. S. 2017, Azamathulla H.M., 2016; Hammouri, N. et al., 2007).

Rainfall-Runoff modeling is one of the most useful applications of GIS. Baiyinbaoligao et al. (2011) has employed GIS for modeling the rainfall-runoff process in the Kuronagi River based on two rainfall stations. Hammouri et al., 2007 applied GIS to model the rainfall-runoff process in the Wadi Madoneh ungauged basin in Jordan. The main objective of this study is to investigate the surface water potential for groundwater artificial recharge. Predicting water quality components is another prominent application of GIS in water management.

Amir Hamzeh et al., 2018 investigates the performance of artificial intelligence techniques including artificial neural network, group method of data handling and support vector machine for predicting water quality components of Tireh River located in the southwest of Iran..

GIS also finds its application in ground water management. In Saro. L et al., 2018's work, groundwater productivity-potential was analyzed using the data mining models of an artificial neural network and a support vector machine in Boryeong city, Korea.

V. CONCLUSION

A Machine learning based GIS for water management is a major area of research to promote a country's economy. The tendency for GIS to become more important in the mainstream of other information technologies means that it will be easier in the future than it has been in the past to synthesize geospatial with temporal water resources data, and to robustly link water resources models to the resulting hydrologic information system. Finally, the great thing that makes GIS attractive in water resources is that in this environment an integrated analysis could be achieved truly. The efficiency of different machine learning-based models has been approved in several studies. But, still a wide research gap is available to make use of GIS and machine learning tools more effectively.

REFERENCES

[1]. Amir H.H, Ali.H.N, Abbas.P (2018),"Water quality prediction using machine learning methods" ,Water Quality Research Journal 53 (1): 3-13.

[2]. Amiri, B. J., and Nakane, K. (2008), "Entire catchment and buffer zone approaches to modeling linkage between river water quality and land cover—A case study of Yamaguchi Prefecture, Japan." Chin. Geogra. Sci., 18(1), 85–92.

[3]. Amiri, B. J., and Nakane, K. (2009), "Comparative prediction of stream water total nitrogen from land cover using artificial neural network and multiple linear regression approaches." Polish J. Environ. Stud., 18(2), 151–160.

[4]. Azamathulla, H. M., Haghiabi, A. H. & Parsaie, A., (2016)," Prediction of side weir discharge coefficient by support vector machine technique.", Water Sci. Technol. Water Supply 16 (4), 1002–1016.

[5]. Baiyinbaoligao, Ding, W., XiangYang, L. (2011), "Application of ArcGIS in the calculation of basins rainfall-runoff.", Procedia Environmental Sciences, 10, 1980–1984.

[6]. Bodeaux M, Prins JG, (2000), "Building land use plans for water master plans: a case study.", ESRI User's Conference, San Diego, CA .

[7]. Chen, Q., Mei, K. & Dahlgren, R. A., (2016), "Impacts of land use and population density on seasonal surface water quality using a modified geographically weighted regression.", Science of the Total Environment **572**, 450.

[8]. Clapcott,J.,Goodwin,E.,&Snelder,T.(2013), "Predictivemodels of benthic macro-invertebrate metrics.", Cawthron Report No. 2301. p. 35.

[9]. Coltin B, McMichael S, Smith T, Fong T. (2016), "Automatic boosted flood mapping from satellite data.", Int J Remote Sens. 37:993–1015.

[10]. Davoodi.M.D., Rezaei.M., Pourghasemi.H.R., Pourtaghie,Z.S.,&Pradhan,B.(2015), "Ground water spring potential mapping using bivariate statistical model and GIS in the Taleghan watershed", Iran. Arabian Journal of Geosciences, 8(2), 913–929.

[11]. De'ath G, Fabricius K. (2010), Water quality as a regional driver of coral biodiversity and macroalgae on the Great Barrier Reef. Ecol Appl. 20:840–850.

[12]. Elith,J.,Leathwick,J.R.,&Hastie,T.(2008).Aworkingguideto boosted regression trees. Journal of Animal Ecology, 77, 802–813.

[13]. Emamgholizadeh, S., Kashi, H., Marofpoor, I. & Zalaghi, E. (2014), "Prediction of water quality parameters of Karoon River (Iran) by artificial intelligence-based models.", Int. J. Environ. Sci. Technol. 11 (3), 645–656.

[14]. Hammouri, N., El-Naqa, A. (2007), "Hydrological modeling of ungauged wadis in arid environments using GIS: a case study of Wadi Madoneh in Jordan.", Revista mexicana de ciencias geológicas, 24, 185–196

[15]. Heddam, S. (2017a), "Generalized regression neural network based approach as a new tool for predicting total dissolved gas (TDG) downstream of spillways of dams: a case study of Columbia River Basin Dams", USA. Environ. Process. 4 (1), 235–253.

[16]. Heddam, S. (2017b), "Multilayer perceptron neural network-based approach for modeling phycocyanin pigment concentrations: case study from lower Charles River buoy", USA. Environ. Sci. Pollut. Res. 23 (17), 17210–17225.

[17]. Heddam, S. (2017c), "Secchi disk depth estimation from water quality parameters: artificial neural network versus multiple linear regression models?", Environ. Process. 3 (2), 525–536.

[18]. Ilayaraja, K., & Ambica, A. (2015). Spatial Distribution of Groundwater Quality Between Injambakkam-Thiruvanmyiur Areas, South East Coast of India. Nature Environment & Pollution Technology, 14(4).

[19]. Jaddi, N. S. & Abdullah, S.,(2017), "A cooperative-competitive masterslave global-best harmony search for ANN optimization and water-quality prediction.", Appl. Soft Comput. 51, 209–224.

[20]. King T.V.; Neilson, B.T.; Rasmussen, M.T., (2018), "Estimating discharge in low-order rivers with high-resolution aerial imagery. Water Resour. Res. 2018, 54, 863–878.

[21]. Koklu, R., Sengorur, B., and Topal, B. (2010). "Water quality assessment using multivariate statistical methods—A case study: Melen River system (Turkey)." Water Resour. Manage., 24(5), 959–978.

[22]. Leathwick J, Elith J, Chadderton W, Rowe D, Hastie T. (2008), "Dispersal, disturbance and the contrasting biogeographies of New Zealand's diadromous and nondiadromous fish species", JBiogeogr. 35:1481–1497.

[23]. Lee M.J, Kang J.E, Kim G. (2015), "Application of fuzzy combination operators to flood vulnerability assessments in Seoul, Korea", Geocarto Int. 30:1052–1075.

[24]. Lee S, Lee CW. (2015), “Application of decision-tree model to groundwater productivity-potential mapping”, Sustainability. 7:13416–13432.
[25]. Lee S. (2014), “Geological Application of Geographic Information System”, Daejeon: International School for Geoscience Resources of Korea Institute of Geoscience and Mineral Resources. p. 143.
[26]. Maidment DR, Jones NL, Strassberg G (2004), “Arc hydro groundwater geodatabase. Center for Research in Water Resources”, University of Texas, Austin, TX.
[27]. Manap, M.A., Nampak, H., Pradhan, B., Lee, S., Soleiman, W.N.A., & Ramli, M.F. (2012), “Application of probabilistic based frequency ratio modeling ground water potential mapping using remote sensing data and GIS”, Arabian Journal of Geosciences.
[28]. Naghibi, S. A., & Pourghasemi, H. R. (2015), “A comparative assessmentbetween three machine learning models and their performance comparison by bivariate and multivariate statistical methods in groundwater potential mapping.”, Water Resources Management.
[29]. Nikoo, M. R. & Mahjouri, N, (2015), “Water quality zoning using probabilistic support vector machines and self-organizing maps”, Water Resour. Manage. 27 (7), 2577–2594.
[30]. Parsaie, A., Haghiabi, A. H., Saneie, M. & Torabi, H, (2018), “Applications of soft computing techniques for prediction of energy dissipation on stepped spillways.”, Neural Comput. Appl.
[31]. Radin HA (2006), “Using ArcGIS for preprocessing and postprocessing a Modflow-96 groundwater model. GIS and Water Resources IV”, AWRA Spring Specialty Conference, May.
[32]. Razandi, Y., Pourghasemi, H.R., Samani Neisani, N., & Rahmati, O. (2015). “Application of analytical hierarchy process, frequency ratio, and certainty factor models for groundwater potential mapping using GIS”, Earth Science Informatics.
[33]. Rindahl B, Bennett R (2013), “Design and implementation of a data-centered groundwater modeling system”, Proceedings ESRI annual users conference. San Diego, CA.
[34]. Rossetto, R. U. D. Y., Borsi, I. A. C. O. P. O., Schifani, C. L. A. U. D. I. O., Bonari, E. N. R. I. C. O., Mogorovich, P. A. O. L. O., & Primicerio, M. A. R. I. O. (2013). SID&GRID: Integrating hydrological modeling in GIS environment. Rendiconti Online Società Geologica Italiana, 24, 282-283.
[35]. Rossman LA (2000), “EPANET 2 users manual. U.S. Environmental Protection Agency.”, Water Supply and Water Resources Division, National Risk Management Research Laboratory, Cincinnati, OH, 45268 Report EPA/600/R-00/057, Sep.
[36]. Saro L, Soo-Min Hong & Hyung-Sup Jung (2018), “GIS-based groundwater potential mapping using artificial neural network and support vector machine models: the case of Boryeong city in Korea, Geocarto”, International, 33:8, 847-861, DOI: 10.1080/10106049.2017.1303091.
[37]. Sener, S., Şener, E. & Davraz, A. (2017), “Evaluation of water quality using water quality index (WQI) method and GIS in Aksu River (SW-Turkey).”, Sci. Total Environ. 584–585, 131–144.
[38]. Shokoohi, M., Tabesh, M., Nazif, S. & Dini, M.,(2017), “Water quality based multi-objective optimal design of water distribution systems.”, Water Resour. Manage. 31 (1), 93–108.
[39]. Swain, N. R., Latu, K., Christensen, S. D., Jones, N. L., Nelson, E. J., Ames, D. P., & Williams, G. P. (2015). A review of open source software solutions for developing water resources web applications. Environmental Modelling & Software, 67, 108-117.
[40]. Wu Z.Y, Wang RH, Diaz D, Walski T (2003), “Mining water consumption and GIS-based data for loading water distribution models.”, ASCE World Water Congress 118:23
[41]. Zhang, W., Wang, Y., Peng, H., Li, Y., Tang, J. & Wu, K. B.,(2010), “A coupled water quantity–quality model for water allocation analysis”, Water Resour. Manage. 24 (3), 485–511.

J.Charles1 “A Suyrvey On Machine Learning Methods In Gis Based Water Modeling” International Journal of Computational Engineering Research (IJCER), vol. 08, no. 08, 2018, pp. 26-30.