Smart survey on recent trends in water level, drought and water quality analysis system

J Balajee1* and M A Saleem Durai2
1School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, Tamilnadu, India.
2Department of IoT, School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamilnadu, India.
Email: balajij2011@vit.ac.in1*; masaleemdurai@vit.ac.in2

Abstract. Over 200 million yearly reports of diseases identified with scarce water and sanitation conditions, 5-10 million deaths occurred worldwide. Water quality checking has subsequently gotten important to supply clean and safe water. This survey work depicts the fundamental explanation behind the requirement for robust and productive Water level, Drought, and water quality control in the level framework, which will keep human assets healthy, sustainable and diminish water use for household purposes. Climate change and variability have so many significant impacts caused by the natural environment's water system. Incredible methods, collection of water samples are tested alone and analyzed in water laboratories. However, it is not always easy to capture, analyze, and rapidly disseminate information to relevant users to make timely and well-informed decisions. The review work encompasses traditional methods based on Machine Learning (ML), and Deep Learning (DL) approaches.

1. Introduction
Due to population growth and economic growth, India faces a significant natural resource problem in water [Ranjan et al. 2020]. Most of the water levels have been contaminated due to excess pollutants, which are mostly human-made. Thus certifying water purity is a significant challenge. With rapid industrialization and the latest technology, greater emphasis on agricultural development, and fertilizers and pesticides, large-scale pollution occurs in the aquatic environment that drives the decline in marine life water quality[4]. Water bodies discharge from sewage and discharge industries due to run-off agricultural sectors, run-off urban areas, floods, droughts including pollution point and non-source, pollution, and lack of educational awareness facility among users.
User involvement in observing aspects such as sanitation, environmental hygiene, storage, and disposal are distinctive elements for maintaining water bodies' quality. The strength of lakes, rivers, and other water bodies and their biological diversification are directly associated with the health of almost every element of the ecosystem. Waterborne diseases such as water contaminated by environmental factors cause death in the environment, and socio-economic progress is slowly spreading. About 5 million people are said to have died because of waterborne diseases worldwide (Indian Water Resources Information System, 2017). The sample of contaminated water is affected by water quality, as shown in 'figure 1', for six reasons: agriculture, wastewater, and sewage, oil pollution, radioactive waste, urban development, and plastics. User involvement in observing aspects such as sanitation, environmental hygiene, storage, and disposal are distinctive elements for maintaining water quality. The strength of lakes, rivers, and other water bodies and their biological diversification are directly associated with the health of almost every element of the ecosystem. Waterborne diseases such as water contaminated by environmental factors cause death in the environment, and socio-economic progress is slowly spreading. About 5 million people are said to have died because of waterborne diseases worldwide (Indian Water Resources Information System, 2017). The sample of contaminated water is affected by water quality, as shown in 'figure 1', for six reasons: agriculture, wastewater, and sewage, oil pollution, radioactive waste, urban development, and plastics.

**Table 1. Ideal water ranges per WHO standards**

| Parameters Monitored | Quality Range | Units |
|----------------------|---------------|-------|
| Turbidity            | 5 to 10       | NTU   |
| pH                   | 6.5 to 8.5    | pH    |
| Conductivity         | 300 to 800    | MicroS/cm |

The primary point of the structure taking care of consistent water quality is to evaluate physical, material properties, and earth quality standards, with a definite objective to identify the water parameters’ types and give an early warning of the dangers. The structure additionally provides a consistent examination of the accumulated information and recommends reasonable medical measures when water pollution slows down. Ideal water ranges following WHO standards are listed in Table 1. This work's objectives were to
survey the activities and quality monitoring systems at the smart water level regarding the application using communication technology and sensors based on a different machine and deep learning methods. The task considers a facility to inform users and officials concerned over the water level's variability.

2. Literature survey

Various technical works on the Assessment of water quality worldwide have presented at the research level from which this survey work referred many papers for study.

2.1. Manual water quality monitoring
Water plays an essential role in all aspects of our lives, and its quality is deteriorating with ever-increasing pollution due to urbanization, industrialization, and population growth. To sustain the quality of life, it is imperative to detect water pollutants, causing contamination. Typically, water quality detection is time-consuming and cumbersome, requiring manual lab analysis and statistical inferences (Gazzaz et al., 2012). Daud et al. (2017), in a research study, collected various water samples and tested for different parameters that were compared against the World Health Organization (WHO) standards and National Environmental Quality Standards (NEQS).

2.2. Machine learning classification methods based on water quality monitoring
Classification is a major problem in machine learning and data mining. It is widely used in many real-world applications. To build an assortment, a user first collects a set of tutorials on training examples/instances of pre-classes. A classification algorithm then uses specialized classification phase training data to assign limited classes to test cases (for evaluation) or future events (application). This section discusses some machine learning classification methods.

Muharemi et al., 2018 proposed the Nearest Neighbor Algorithm (NN) and the Neural Network of Classification dependent on Logistic Regression to acquire a sufficient answer for address changes like drinking water. Haghiaib et al. (2018) researched the presentation of artificial insight strategies that incorporate the Artificial Neural Network (ANN), the Group Data Management Method (GDMM), and the Support Vector Machine (SVM) to foresee the parts of the water nature of the Tireh River situated in southwestern Iran. During the improvement cycle of ANN and SVM, it was discovered that tanSig and Radial Bias Function (RBF) as move and center functions have the best exhibition among the tried functions (Haghiaib et al., 2018). Chou et al. (2018) led an examination to decide the water quality in the repository. Four notable artificial insight strategies, Artificial Neural Networks (ANN), support vector machines, classification and regression trees, and linear regression were utilized to dissect reference situations and sets. A simple to-utilize interface was built up that incorporates a metaheuristic regression model to assess prescient execution and contrast it with those of the two constituent situations.

The ANN model was more precise than the other extraordinary models, sets, and meta-heuristic regression hybrids. (Chou et al., 2018), proposed another anomaly recognition algorithm for water quality information utilizing double development windows after continuously recognizing chronicled design oddity information. The algorithm depends on measurable models, autoregressive straight mix model. The algorithm has tried utilizing water quality PH information three months from a genuine water quality checking station on a waterway framework.

Mohammad et al. (2015) explored water quality issues, utilizing three unique algorithms, SVM, and two artificial neural network strategies. The exhibition looks at the use of R squared (R 2), Root Means Square Error (RMSE) and Means Absolute Error (MAE). The result they completed, the support vector machine algorithm, is a severe neural network. This work takes us to reconstruct SVM's neural network and Muharemi et al. (2018a) because it promises better results. The best result is to realize the use of artificial neural networks and non-direct autoregression.
2.3. Deep learning methods based water level, drought and water quality monitoring

Deep Learning is a subfield of Machine Learning where a since quite a long-known algorithm a Neural Network (NN) or an Artificial Neural Network (ANN) or a Deep Neural Network (DNN) or a Feed-Forward Neural Network (FFNN), alluding to a similar algorithm, is utilized to plan highlights into input, hidden and output layer. Framed by an input layer, hidden layers, and an output layer, ANNs present an effective method to learn straight and non-direct connections among input and output sets appeared in 'figure 2'. This section discusses the operation of different deep learning methods based water quality analysis system.

Convolutional Neural Networks (CNN), which are developed in profound various leveled models and are fit for removing inherent features shown in 'figure 3', have been used in water quality classification [Fan Hu et al. 2015] and accomplishing excellent outcomes [Gui-Song Xia et al. 2017]. NNs (Goodfellow et al. 2014) consists of two seemingly separate CNNs working in union and competing in a min-max game. NNs are initially used as generative models, randomly generating new samples from a dataset to appear as if they are from the originating dataset when visualized (Gautam et al. 2020).

Recurrent Neural Networks (RNN) aim to keep the information gained from earlier parts of a data sample in the memory and move it to the later parts of the same data sample to ensure better knowledge discovery (Goodfellow et al., 2016). A simple RNN implementation lacks the practicality in long
sequences, such as long paragraphs, as it is common to encounter the vanishing gradient problem while training. More complex RNN implementations like Long Short-Term Memory (LSTM) (Hochreiter & Schmidhuber 2019) Networks or Gated Recurrent Unit (GRU) Networks (Cho et al. 2014) solve this problem but have greater computational complexity.

![Figure 4. Computation-wise comparison of RNN, LSTM, and GRU nodes](image)

However, data-driven models are increasingly being used in a variety of water quality applications. One such application area is predicting surface water quality. Li et al. (2019) propose an ensemble approach that combines three RNN models with Dempster/Shafer (DS) evidence theory to predict water quality. The results of three RNN models, namely LSTM, GRU, and Elman Neural Network, are combined by DS evidence theory for the final output are shown in 'figure 4'. The combined model predicts at most 50 hours in advance, and the results show that the model accuracy reduces significantly over 25 hours.

Liu et al. (2019) use an LSTM model to forecast drinking water quality for up to 6 months. Zou et al. (2020) develop an ensemble approach to predict water quality data, such as pH, DO, CODMn, and NH3-N. The approach is based on using three LSTM models that different size interval data feed each of them, and the final prediction is a combination of three LSTM models result. Banerjee et al. (2019) choose the indicators, namely dissolved oxygen and zooplankton abundance, to reflect the water quality level.

2.4. Water level, drought and water quality monitoring based on wireless sensor network

Wireless sensor networks have generally been received to monitor various wonders in the climate [M. Wu et.al 2016]. For instance, WSNs are utilized to monitor air quality [Guanochanga et al., 2018] and water quality [Adu-Manu et al., 2018]. The appropriation of WSNs permits the end of issues related to traditional observing methodologies. Vijayakumar and Ramya (2015) planned and built up an ease, ongoing water quality observing framework utilizing the Internet of Things (IoT) advancements[28]. Nasser, Ali, Karim, and Belhaoauri (2013) planned the actual use of energy efficiency, self-configuration, and reusable based on the WSN-water quality observation framework. The ZigBee sensor hub's proposed system that opens the device includes (Squidbee), a server farm associated with the Internet, and includes a data entry channel and a fixed client station for network workers. Lambrou (2014) planned and established a minimal workload in-pipe sensor network architecture to observe drinking water quality at the user's destination.

2.5. Gaps in literature
The majority of pre-existing techniques have certain conditions and issues because it points out some of them are ignored:

1. Water level, Drought, and water quality recognition using feature selection techniques integration can further increase accuracy.
2. Most of the existing techniques should be as far as the essential aspects of Water level, Drought, and water quality are concerned.
3. The integration of feature selection techniques and classification mechanisms has been neglected to improve the accuracy ratio to recognize Water level, Drought, and water quality.

3. Conclusion
This survey work has presented Water level, Drought, and water quality monitoring using Machine Learning, Deep Learning, and Smart Wireless Sensor Networks. An extensive group of information exists on data mining utilizing conventional ML procedures and late utilizing Deep Learning models inferable from renewed enthusiasm for Water level, Drought, and water quality examination. This review shows that the utilization of deep learning models for Water level, Drought, and water quality testing is still in the early stages, as documented in the literature. Moreover, the Long Short Term Memory (LSTM) and Gated Recurrent Unit (GRU) application for Water level, Drought, and water quality examination supposedly dependent on the right now investigated literature are non-existent. Therefore, the energetic exploration direction will be incorporated into deep learning technology to inspect further Water level, Drought and water quality, and water level analysis.

References

[1] Fan Hu, Gui-Song Xia, Jingwen Hu and Liangpei Zhang 2015 Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery Remote Sensing 7(11), pp 14680–707.
[2] Gui-Song Xia, Jingwen Hu, Fan Hu, Baoguang Shi, Xiang Bai, Yanfei Zhong, Liangpei Zhang and Xiaoliang Lu 2017 Aid: A benchmark data set for performance evaluation of aerial scene classification IEEE Transactions on Geoscience and Remote Sensing 55(7), pp 3965–81.
[3] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A and Bengio Y 2014 Generative adversarial nets Advances in Neural Information Processing Systems pp 2672–80.
[4] Gautam A, Sit M and Demir I 2020 Realistic river image synthesis using deep generative adversarial networks arXiv preprint arXiv:2003.00826.
[5] Goodfellow I, Bengio Y and Courville A 2016 Deep Learning. MIT press.
[6] Hochreiter S and Nearing G S 2019 Toward improved predictions in ungauged basins: exploiting the power of machine learning Water Resources Research.
[7] Cho K, Van Merriënboer B, Bahdanau D and Bengio Y 2014 On the properties of neural machine translation: encoder-decoder approaches arXiv preprint arXiv:1409.1259.
[8] Li L, Jiang P, Xu H, Lin G, Guo D and Wu H 2019 Water quality prediction based on recurrent neural network and improved evidence theory: a case study of Qiantang River. China. Environmental Science and Pollution Research 26 (19), 19879–96.
[9] Liu Y, Qin H, Zhang Z, Yao L, Wang Y, Li J, Liu G and Zhou J 2019 Deriving reservoir operation rule based on Bayesian deep learning method considering multiple uncertainties Journal of Hydrology 579, 124207.
[10] Zou Q, Xiong Q, Li Q, Yi H, Yu Y and Wu C 2020 A water quality prediction method based on the multi-time scale bidirectional long short-term memory network Environmental Science and Pollution Research pp 1–12.
[11] Banerjee A, Chakrabarty M, Rakshit N, Bhowmick A R and Ray S 2019 Environmental factors as indicators of dissolved oxygen concentration and zooplankton abundance: deep learning versus traditional regression approach Ecological Indicators 100, pp 99–117.

[12] Muharemi F, Logofatu D, Andersson C and Leon F 2018 Approaches to building a detection model for water quality: A case study Modern Approaches for Intelligent Information and Database Systems pp 173–83.

[13] Haghiani A H, Nasrolahi A H and Parsaie A 2018 Water quality prediction using machine learning methods Water Quality Research Journal 53(1), pp 3-13.

[14] Vinoth Kumar V, Ramamooorthy S, Philip Kumar V, Prabu M, Balajee J M 2021 Design and evaluation of Wi-Fi offloading mechanism in heterogeneous networks International Journal of e-Collaboration 17(1), pp 62-70.

[15] Chou J S, Ho C C and Hoang H S 2018 Determining quality of water in reservoir using machine learning Ecological informatics 44, pp 57-75.

[16] Mohammadpour R, Shaharuddin S, Chang C K, Zakaria N A, Ab Ghani A and Chan N W 2015 Prediction of water quality index in constructed wetlands using support vector machine Environmental Science and Pollution Research 22(8), pp 6208–19.

[17] Muharemi F, Logofatu D and Leon F 2018 Review on general techniques and packages for data imputation in r on a real-world dataset Springer.

[18] Nguyen M and Logofatu D 2018 Applying tree ensemble to detect anomalies in real-world water composition dataset International Conference on Intelligent Data Engineering and Automated Learning pp 429–38.

[19] Gazzaz N M, Yusoff M K, Aris A Z, Juahir H, Ramli M F 2012 Artificial neural network modeling of the water quality index for Kinta River (Malaysia) using water quality variables as predictors Mar. Pollut. Bull. 64, pp 2409–20.

[20] Khan Daud, Nafees Muhammad, Ali Shafaqat, Rizwan, Muhammad, Bajwa Raees, Shakoor Muhammad, Arshad Muhammad, Chatha Shahzad Ali Shahid, Deeba Farah, Murad Waheed, Malook Ijaz and Zhu Shui 2017 Drinking-water quality status and contamination in pakistan-a review BioMed Research International.

[21] Ranjan Aalok 2020 Population and Sustainable Development in India. 10.1007/978-981-32-9212-3.

[22] Wu M, Tan L and Xiong N 2016 Data prediction, compression, and recovery in clustered wireless sensor networks for environmental monitoring applications Information Sciences 329, pp. 800–818.

[23] Guanochanga B, Cachipuendo R, Fuertes W et al. Real-time air pollution monitoring systems using wireless sensor networks connected in a cloud-computing, wrapped up web services Proceedings of the Future Technologies Conference pp. 171–84.

[24] Adu-Manu K S, Tapparello C, Heinzelman W, Katsriku F A and Abdulai J D Water quality monitoring using wireless sensor networks: current trends and future research directions ACM Transactions on Sensor Networks 13(1) p. 4.

[25] Vijayakumar N and Ramya R 2015 The real time monitoring of water quality in the IoT environment IEEE sponsored 2nd international conference on innovations in information, embedded and communication systems(ICIECS), pp 1–4.

[26] Ali N, Karim A L and Belhaouari S 2013 An efficient WirelessSensor Network-based water quality monitoring system ACS inter-national conference on computer systems and applications pp.1–4.

[27] Lambrou T P, Anastasiou C C, Papayioutou C G and Polycarpou M M 2014 A low-cost sensor network for real-time monitoring and contamination detection in drinking water distribution systems IEEE Sensors Journal 14(8), pp 2765–72.

[28] Parveen SH, Shrivastava, Nirvishi, Dominic, Durai, Nalini N, Balajee J M 2019 Comparison of machine learning algorithms to build optimized network intrusion detection system Journal of Computational and Theoretical Nanoscience 16(5-6), 2541-49.