Prediction of Drinking Water Quality for Effective Public Health Management to Avoid Water Borne Diseases

Rajvir Saini[1], Vivek Shrivastava[2]

[1]*Assistant Professor, Department of Management, Kalinga University, Raipur, India. Email: ku.rajvirsaini@kalingauniversity.ac.in ORCID: 0009-0000-6644-0795*

[2]*Assistant Professor, Department of Management, Kalinga University, Raipur, India. Email: ku.vivekshrivastava@kalingauniversity.ac.in ORCID: 0009-0000-9623-4968*

KEYWORDS

Water quality, Drinking water, Public health, Management, Water-borne diseases, Convolutional Neural Network

ABSTRACT

The issue of water quality failure is a persistent global concern that leads to sickness, toxicity, the spread of diseases, and loss of life in metropolitan areas. Various physical, functional, and ecological factors may affect drinking water quality. These concerns include the incursion of pollutants into water pipes, leakage, the formation of disinfectant waste products, the penetration of chemicals or microbes, and pollution. Water-borne Diseases (WD) have emerged as a significant global public health issue. It is imperative to forecast the quality of drinkable water by utilizing an automated model that provides a preventive strategy that may be created to enhance public health management (PHM). This study forecasted the Drinking Water Quality (DWQ) to promote Public Health Management (PHM) strategies to prevent WD. Convolutional Neural Network (CNN) with Inverse Distance Loaded (IDL) interpolation has been used to forecast DWQ. One hundred five water specimens from the research region underwent thorough examination via physiochemical and microbiological analyses. The tested characteristics were pH, turbidity, temperature, fluoride concentration levels, and bacterial counts (Faecal Coliforms) to prevent WD. The test findings indicated that the CNN-IDL projected 35% of water specimens from the domestic pipeline network had bacteria levels below what the WHO considers acceptable. Also, about 57% of the samples from the home pipeline network met the fluoride reference values set by the WHO. The sanitary assessment rating clearly showed that the lack of appropriate measures to protect against multiple barriers resulted in declining water quality for users. The findings of this research will be valuable to administrators in implementing essential strategies to avoid WD before it significantly impacts public health.

1. Introduction

Due to the ongoing population expansion and fast economic growth, there has been an increasing need for pure and potable water, resulting in an acute water shortage [1]-[5]. Specifically, excessive withdrawal of groundwater and the discharge of wastewater harm the quality and accessibility of groundwater resources [9]. Chemical contaminants emitted by human conduct, such as agriculture and manufacturing, are responsible for water pollution issues that directly endanger or potentially jeopardize human health [2]. According to the WHO (2011), water, acting as a transport of pollutants, may cause 72% of illnesses, and globally, 25% of malignancies are linked to water pollution [3]. Continuous exposure to nitrate-contaminated water poses a significant danger to clients, especially newborns, who are more susceptible to the "blue baby phenomenon" resulting from access to nitrate-contaminated drinking water [19]. Consumers play a vital role in promoting the growth of any business, protecting environment and promoting social infrastructure [4]-[16]. Water Distribution Networks (WDN) face several physical, ecological, and practical limitations, which make it challenging to assess the likelihood of DWQ degradation and evaluate the state of pipes [7]. This is because WDN consists of long pipes made of various times and substances, with insufficient and unclear historical information related to their physical and operational efficiency [20]. The complexities of identifying and evaluating the probability of failure are due to the significant time gap between the onset of breakdown and the detection of its consequence [8]. Even more concerning is that investigations into the cause of WD outbreaks or other pollution issues are not conducted until after the events [6]. The degradation of DWQ, along with the intrinsic intricacies of water systems, demonstrates that risk assessment is a very intricate process that may significantly impact societies and public institutions. ANNs have been used in several research projects focused on water distribution systems and pipes. Their primary purpose is to evaluate the likelihood and speed of pipe breakage and leakages [17].

WD results from pathogens in water, leading to many forms of diarrheal illnesses, including cholera and dysentery. Diarrhea is the leading cause of mortality in children, responsible for around 8.5% of all deaths among kids under the age of 5. This results in over 11500 pediatric deaths globally [10]. Approximately 65% of worldwide fatalities caused by diarrhea are attributed to the use of contaminated drinking water and poor personal, home, and communal hygiene practices, as well as inadequate sanitation. Research indicates that WD is responsible for more than 2.3 million fatalities per year on a global scale, and they also result in significant financial consequences of approximately 13 billion USD globally [11]. While most DWQ issues linked to health are caused by bacterial contamination, it is important to note that chemical pollution of water sources may also lead to significant health issues [21]. An instance of this is when water contains nitrate and nitrite, which may occur due to the overuse of fertilizers or effluent infiltration into drinking water and groundwater [12]. Fluorosis is a prevalent issue caused by exposure to elevated quantities of naturally present fluoride in freshwater [18]. This contact may cause discoloration of children's teeth, skeletal fluorosis, and severe physical impairment. Ahmad et al. [13] conducted a study to examine and create a map of the regional distribution of the criteria that determine the DWQ in Peshawar, Pakistan. A total of 108 water specimens, gathered from various locations across the research region, underwent thorough examination via physiochemical and biological analysis. The results indicated that roughly 20% of the region was found to have Fecal Coliforms (FeC) requiring therapy, while around 33% of the area had fluoride levels that also required treatment.

Castro et al. undertook an extensive investigation to investigate the spatial and temporal changes in DWQ mechanisms for the water specimens obtained from Yucatan, Mexico [14]. The outcomes of the classification hierarchy indicate that the spatio-temporal variation of the groundwater gathered in the research is primarily influenced by variables such as the elevated concentration of sulfates in the groundwater, incursion of saltwater, relations among minerals and water, and other human activities [15]. Although earlier research has been conducted, CNNs have hardly been utilized to predict and assess the danger of DWQ issues. Maintaining the DWQ is an essential aspect of PHM. Water that is contaminated may result in the spread of WD, including diarrhea, vomiting, and typhoid, which are major contributors to sickness and death on a global scale. Utilizing sophisticated analytical methods and technology to forecast DWQ may significantly contribute to disease prevention and PHM. The primary objective of the ongoing project is to provide an innovative approach for estimating and forecasting the DWQ in WDNs using CNN for prognostics and PHM.

2. Methodology

A comprehensive study was conducted on 105 water samples collected from various locations within the research zone. The analysis included both physiochemical and microbiological examinations. The examined variables include pH, turbidity, temperature, fluoride levels, and the number of bacteria (specifically, FeC) to mitigate the risk of WD. The parameters pH, turbidity, and temperature were analyzed due to their importance in the chlorination process of drinking water, whereas FeC indicates the existence or lack of harmful bacteria. FeC in the aquatic atmosphere indicates water contamination by animal and human excrement. The DWQ testing and modeling using the IDL interpolation approach. Many studies have routinely used the IDL approach because of its resilience. The research region employed an unknown location to anticipate the DWQ values by measuring the features of the water. The predictions made by the CNN technique also rely on a linear array of loads assigned to known points to estimate the values of undetermined points. The mathematical formulation for IDL is represented by Equation (1):

$$Y_x = \frac{\sum_{i=1}^{n} \frac{y_i}{d_i^p}}{\sum_{i=1}^{n} \frac{1}{d_i^p}} \tag{1}$$

Y_x represents the amount present at an undetermined location. Let x represent the levels at the observed sites i. The distance from x to i is denoted as d_i, and the magnitude by which the distance is scaled is represented by p. The output of IDL has been fed to the input layer of CNN to predict DWQ for effective PHM to prevent WD. A CNN's design includes important parts like Fully Connected (FC) layers, Convolutional Layers (CL) layers, and Pooling Layers (PL) layers. Each layer in the system is very important for pulling and abstracting characteristics. A key part of CNNs is CL, and the main operation is the convolution operation. During the process, a kernel filter is applied to the input data to get specific features related to DWQ. The convolution process can be expressed by Equation (2):

$$S(i,j) = (I * K)(i,j) = \sum m \sum n I(m,n) K(i-m, j-n) \quad (2)$$

Here, $S(i,j)$ is the output of the convolution operation, I is the input image, and K is the filter.

Activation Function: In normal situations, a non-linear activation function like the Rectified Linear Unit (ReLU) is used after the convolution process. This is how the ReLU function is described as follows:

$$f(x) = max(0,x) \quad (3)$$

PL: The PL function lowers the input's spatial dimensions, which makes it easier to compute and better at maintaining translational invariance. Max pooling is a type of pooling that is often used in computer vision. This operation has been designated as:

$$O(i,j) = \max \begin{bmatrix} I(2i,2j) & I(2i+1,2j) \\ I(2i,2j+1) & I(2i+1,2j+1) \end{bmatrix} \quad (4)$$

FC layers: The network usually ends with one or more FC levels once there are enough CL and PL layers. The links between neurons in each layer create a complete pattern of connectivity in which every neuron is related to every neuron in both the layers above and below it.

Using CNNs to predict DWQ provides a powerful tool for PHM. By leveraging Deep Learning (DL), it is possible to detect and respond to water quality issues more quickly and accurately, thereby preventing WD and safeguarding public health.

3. Results and discussion

FeC found in excrement was used as an indicator to check if drinking water sources were contaminated with bacteria. Fig. 1 shows that tubewell and borewell water were both healthy. However, 35% of water specimens from the domestic pipeline network had bacteria levels below what the WHO considers acceptable. In the same way, about 57% of the samples from the home pipeline network met the fluoride reference values set by the WHO. Regarding fluoride content, borewells were the next source that worsened, with 64% of the specimens being within the WHO recommended value.

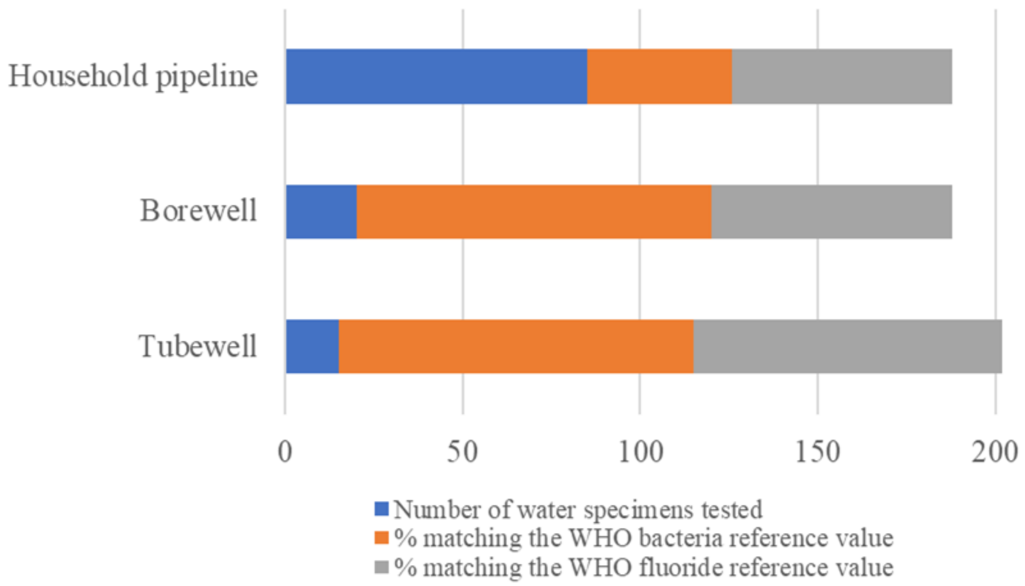

Fig. 1 Testing results of DWQ specimens using CNN-IDL framework

Fig. 2 shows the performance analysis of various DL models for predicting DWQ to promote PHM strategies to prevent WD. The CNN-IDL model has almost perfect prediction abilities, with an accuracy of 0.997, precision of 0.996, recall of 1, and F1-score of 0.9965. The CNN model has 0.983 accuracy, 0.98 precision, 0.987 recall, and 0.983 F1-score. These data prove the CNN model predicts

DWQ well. Despite a slight decrease in accuracy, precision, recall, and F1-score, the RNN model performs well. With 0.987 accuracy, 0.983 precision, 0.987 recall, and 0.982 F1-score, the LSTM model performs well. The most successful model, CNN-IDL, predicts DWQ, providing essential information for efficient PHM and the prevention of WD.

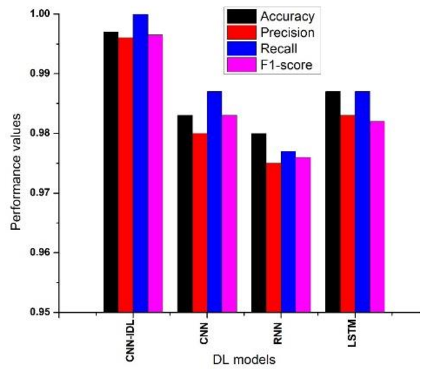

Figure 2. Performance analysis of various DL models for predicting DWQ to promote PHM strategies to prevent WD

4. Conclusion and future scope

WD has become a major public health problem around the world. It is important to predict drinking water quality using an automatic model that offers a preventative approach. This model could be used to improve PHM. The study predicted that the DWQ would help PHM devise ways to stop WD. Convolutional Neural Network (CNN) with Inverse Distance Loaded (IDL) approximation has been used to predict the DWQ. One hundred-five water samples collected from all over the study area were carefully examined using physiochemical and microbiological methods. PH, turbidity, temperature, fluoride concentration levels, and bacterial numbers (FeC) were all checked to prevent WD. The test results showed that 35% of the water samples from the residential piping network had bacteria levels below what the WHO believes safe. In the same way, the WHO set fluoride standard levels that were met by about 57% of the samples from the home piping network.

Reference

[1] Su, F., Wu, J., & He, S. (2019). Set pair analysis-Markov chain model for groundwater quality assessment and prediction: A case study of Xi'an city, China. *Human and Ecological Risk Assessment: An International Journal*, *25*(1-2), 158-175.

[2] Tasneem, A., Sarker, P., Akter, S., Mouna, S. S. P., Rahaman, M. S., Mohinuzzaman, M., ... & Kabir, M. M. (2021). Textile wastewater treatment by combination of chemical and phytoremediation processes. *Pollution*, *7*(1), 43-54.

[3] Jayapriya, R. (2021). Scientometrics Analysis on Water Treatment During 2011 to 2020. Indian Journal of Information Sources and Services, 11(2), 58–63.

[4] Raman, A. & Ramachandaran, S. D. (2023). Factor Influencing Consumer's Adoption of Electric Car in Malaysia. TEM Journal. Volume 12, Issue 4, pp 2603-2612, ISSN 2217-8309, DOI: 10.18421/TEM124-50.

[5] Raman, A., Suhartanto, D., & Shaharun, M.H.B. Delightful Customer Experience: An Antecedent for Profitability and Sustainable Growth of Airline Businesses. Preprints (2023), 2023121838. https://doi.org/10.20944/preprints202312.1838.v1

[6] Stephen, K. V. K., Mathivanan, V., Manalang, A. R., Udinookkaran, P., De Vera, R. P. N., Shaikh, M. T., & Al-Harthy, F. R. A. (2023). IOT-Based Generic Health Monitoring with Cardiac Classification Using Edge Computing. Journal of Internet Services and Information Security, 13(2), 128-145.

[7] Malek Mohammadi, M., Najafi, M., Kermanshachi, S., Kaushal, V., & Serajiantehrani, R. (2020). Factors influencing the condition of sewer pipes: State-of-the-art review. *Journal of Pipeline Systems Engineering and Practice*, *11*(4),

03120002.

[8] Aghapoor Khameneh, P., Miri Lavasani, S. M., Nabizadeh Nodehi, R., & Arjmandi, R. (2020). Water distribution network failure analysis under uncertainty. *International Journal of Environmental Science and Technology*, *17*, 421-432.

[9] Robles, T., Alcarria, R., De Andrés, D.M., De la Cruz, M.N., Calero, R., Iglesias, S., & Lopez, M. (2015). An IoT based reference architecture for smart water management processes. Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 6(1), 4-23.

[10] Day, A. S. (2020). Chronic Diarrhoea in Infants and Children: Approaching and Managing the Problem. *Current Treatment Options in Pediatrics*, *6*, 1-11.

[11] Oon, Y. L., Oon, Y. S., Ayaz, M., Deng, M., Li, L., & Song, K. (2023). Water-borne pathogens detection technologies: advances, challenges, and future perspectives. *Frontiers in Microbiology*, *14*, 1286923.

[12] Arora, G. (2024). Desing of VLSI Architecture for a flexible testbed of Artificial Neural Network for training and testing on FPGA. Journal of VLSI Circuits and Systems, 6(1), 30-35.

[13] Ahmad, M., Jamal, A., Tang, X. W., Al-Sughaiyer, M. A., Al-Ahmadi, H. M., & Ahmad, F. (2020). Assessing potable water quality and identifying areas of water-borne diarrheal and fluorosis health risks using spatial interpolation in Peshawar, Pakistan. *Water*, *12*(8), 2163.

[14] Srkalović, D., & Stjepić Srkalović, Ž. (2020). Groundwater Vulnerability determination of Northeastern Bosnia According to GLA Method. Archives for Technical Sciences, 2(23), 9–20.

[15] Li, P., & Wu, J. (2019). Drinking water quality and public health. Exposure and Health, 11(2), 73-79.

[16] Prashanth, B. Arasu, R. & Karunanithy, D. (2024). Perceptual Study on Higher Level Digitilization Among Managers in the Logistics Industry. The Journal of Distribution Science, 22(1), 25-36, https://doi.org/10.15722/jds.22.01.202401.25.

[17] Dawood, T., Elwakil, E., Novoa, H. M., & Gárate Delgado, J. F. (2020). Water pipe failure prediction and risk models: state-of-the-art review. Canadian Journal of Civil Engineering, 47(10), 1117-1127.

[18] Thooyamani K.P., et.al Random projection based data perturbation using geometric transformation, World Applied Sciences Journal, V-29, I-14, PP:19-24, 2014.

[19] Pacheco Castro, R., Pacheco Ávila, J., Ye, M., & Cabrera Sansores, A. (2018). Groundwater quality: analysis of its temporal and spatial variability in a karst aquifer. Groundwater, 56(1), 62-72.

[20] Okan, A., & Christian, C. (2024). Capture of a New-born Shortfin Mako Shark Isurus Oxyrinchus (Lamniformes: Lamnidae), with Updated Records from the Turkish Marine Waters. Natural and Engineering Sciences, 9(1), 01-09.

[21] Madhav, S., Ahamad, A., Singh, A. K., Kushawaha, J., Chauhan, J. S., Sharma, S., & Singh, P. (2020). Water pollutants: sources and impact on the environment and human health. Sensors in water pollutants monitoring: Role of material, 43-62