Data-Driven Techniques and Data Analytics in Water Treatment Facilities: Innovative Safety Protocols and Optimization

.

ABSTRACT

Aim: To examine harnessing of data-driven techniques and data analytics in water treatment facilities considering innovative safety protocols and process optimization.

Problem Statement: The most significant part of human existence and industrial operations has been linked to water which often becomes vulnerable to hazardous contaminants which are brought on through natural processes and human activity.

Significance of Study: The use of water treatment facilities in order to make water a sustainable natural resource for the whole world is vital. To achieve this, it is imperative to monitor and classify water quality.

Methodology: Recent literature materials in form of books, journals and relevant published articles in the area of data-driven techniques and data analytics in water treatment facilities were consulted.

Discussion: This review article has critically examined harnessing of data-driven and data analytics in water treatment facilities with consideration also given to innovative safety protocols and process optimization. The basic fundamental principle of water treatment facilities were discussed alongside the application of data-driven techniques. Safety protocols in water treatment facilities were discussed. Data driven techniques in water treatment facilities and applications of machine learning and artificial intelligence techniques to water treatment were explained. It was noticed that AI-Driven solutions transforms real-time water quality analysis and offers a proactive approach to water treatment. The integration of machine learning and AI algorithms capabilities in water treatment facilities can now respond quickly to fluctuations and continuously monitor key parameters fluctuations in water quality. Consideration was given to artificial intelligence as data analytics optimization tool for water treatment facilities. Future prospects, limitations and challenges of the study were stated.

Conclusion:AI algorithms and historical data can be harnessed into water treatment facilities in order to assess performance patterns, predict equipment malfunctions and proactively plan the maintenance activities.

Keywords: Data-Driven Techniques, Water Treatment Facilities, Safety Protocols, Process Optimization, Data Analytics

1. INTRODUCTION

Water has been stated to be the most significant part of industrial operations and human existence. However, It is presentlyextremelyvulnerable tohazardous contaminants brought on via both natural processes and human activity. Accessibility to water in healthy and safe manners is a little bit challenging around the globe. Thus, it is imperative to classify and monitor the quality of water, however, the major problem is that, with the present technology, sufficient parametric quality metrics are not available. Aquatic systems' activities (aquaponics, aquaculture and hydroponics) have improved resulting from the increase in human population. As a result of this, the main nutrient load, which are majorly phosphorus and nitrogen, drained to bodies of water bodies has greatly increased causing havoc in different water habitats. Therefore, it is essential to develop and improve water treatment facilities in a sustainable manner with the aid of modern technologies [1-3].

Water has been noticed to be an exceptional resource because of its necessity on daily basis. However, the world is facing the global water crisis challenges characterized by pollution, scarcity and ineffective management techniques, which requires a thoughtful revolution in the status quo [2]. The persistent burden on water resources has influenced for a paradigm shift in techniques of handling water conservation, treatment and management. Several limitations have been attached to the traditional methods which were once assumed to be adequate enough in the optimization of resource usage, provision of real-time insights and mitigation of the water pollution with reference to its environmental impacts. The conventional practices inadequacies in addressing the complications of water-related cases has called for the cutting-edge technologies exploration having dependable revolutionized ways via which man can preserve and interact with this finite resource [4].

Facilities of wastewater- and water- treatment alongside numerous biological and industrial systems that are functions of various resources must have access to clean and sustainable water. Treatment facilities must be able to handle complex regulatory steps to fulfill the improving quality standards, in addition to enhancing infrastructure and providing customer wants for quality of life. Around 300–400 million tons of contaminants have been reported to be discharged into global water on yearly basis causing water pollution, which is a huge burden on water quality management. This is only complex based on the fact that some nations remain having severe polluted rivers, which alter terrestrial and aquatic life together with human life. These problems are steadilydeteriorating as nations continue to modernize and industrialize [5]. Improvement of water-related applications has become subject of discussion by researchers around the globe. Few years back, the major focus has being on modeling and developing optimal, intelligent and economical models in order to help in resolving this great challenge [6].

Nonetheless, water treatment facilities require some safety protocols for effective plant management which is attributed with some difficulties that usually have no easy resolutions.Water treatment characteristically involves a complicated and decentralized process. Equally, industries across the globe usually have challenges in finding quality employees. That lack of experienced operators and labor shortage has caused devastating and potentially costly human errors. With reference to labor and water treatment facility management, the development of a water treatment schedule incorporated with safety aids the minimization of the risk of mistakes and injuries. The creation and adherence to a maintenance schedule in water treatment facilities creates room for potential areas of importance to get addressed early [7]. Equipment maintenance also improves its functionality and lifespan which helps in avoiding early replacement costs. It is important that safety should also take the lead in this regard because many employees have to work at distanced heights and handle moving equipment. This is an established fact that safety protocols reduce the potential of an accident in water treatment facilities. Thus, there is need to incorporate artificial intelligence and data analytics in the optimization process for effective and dependable outputs [8].

In the previous years, Artificial Intelligence (AI) has arisen as aninspirational tool with heightened hope serving as a powerful catalyst to reshape industries across the globe. AI-Driven solutions stand as a transformative route to unravel novel opportunities in water treatment facilities and management. The arrival of AI technologies has allowed humans to analyze and process vast datasets which empower humans in having broadened understanding of complicated hydrological systems and in making informed decisions. Viathe harnessing of the power of AI-Driven solutions, humans have the opportunities of unlocking unprecedented potentials to address the complex challenges evolving around water resources having unparalleled efficiency and precision which usher in an advanced period of water sustainability and security. Invaluable insights and optimization of water management techniques can be unlocked via the amalgamation of AI's computational competency [9-11]. Currently, researchers are embarking on a route of insight and innovation which will facilitate humans towards data-driven, optimized and sustainable water treatment and facility management solutions via the convergence of AI's computational expertise.

Figure 1 represents various applicable models that can be utilized in achieving this targets and objectives. These include the recurrent neural network (RNN), the artificial neural network (ANN), fuzzy neural network (FNN), deep neural network (DNN)and convolutional neural network (CNN). The recurrent neural network is a type of neural network which operates via iteration in the path of sequence evolution while the sequence data are taken as the input data. RNN has parameter sharing, memory and completeness of Turing which makes it beneficial for learning of non-linear types in time-series problems. ANNs are mathematical models which simulate biological neural networks behavior in data processing. They make use of unit nodes to mimic neurons and execute information processing via the adjustment of interconnected weights among numerous nodes within the network. FNN is a hybrid NN model which merges fuzzy logic and ANNs in handling ambiguous or uncertain situations. It adopts fuzzy logic reasoning in processing input data and then uses an ANN in training and producing the results. DNN is a form of ANN that comprises numerous hidden layers placed between the output and input layers. The DNN deep architecture enables it to learn ordereddata representations via the combination of lower-level features in successive layers to learn higher-level characteristics. CNNs are a kind of deep learning approach that uses a deep structure feedforward neural network and convolutional computations. CNNs are broadly used in natural language processing, computer vision and other fields. CNNs adopt convolutional layers in the extraction of intricate characteristics from input images, pooling of layers in order to reduce the property of dimensionality, and fully related layers for categorization tasks.

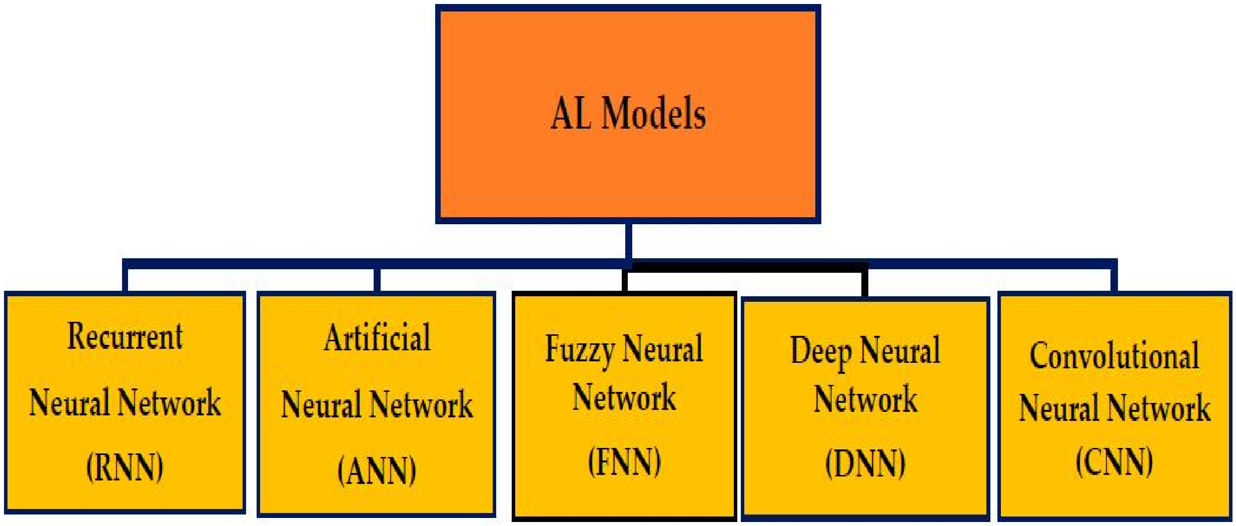

Figure 1: Applicable AI models for data driven and analytics optimization of water treatment facilities

Historically, the traditional and conventional water treatment techniques have been the root and foundation for the provision of drinkable and safe water to communities around the globe using dependable water treatment facilities. These techniques, which are coagulation, sedimentation, flocculation, filtration, adsorption and disinfection, have obviously played a vital role in combating transmittable diseases via water in order to ensure the general public well-being [12]. However, they have their respective attributed problems and shortcomings. The traditional water purifying techniques and water treatment facilities monitoring are usually dependable on manual monitoring and fixed schedules which result potential delays and inefficiencies in the detection of issues that have to do with water quality. Nonetheless, these techniques effectiveness can be affected by different properties which are often found in various water sources and thus make the optimization strategies of the water treatment facilities management to be tasking and difficult [13].

The initiation of AI-Driven solutions transforms real-time water quality analysis and monitoring and thus offers a proactive approach to water treatment. The integration of machine learning and AI algorithms capabilities in water treatment facilities can now respond quickly to fluctuations and continuously monitor key parameters fluctuations in water quality. AIDriven and data analytics systems can handle wide volumes of data from different satellites, sensors and various monitoring devices which provide realtime and accurate insights into water quality patterns and potential contamination cases [14]. AI-data driven natural language processing proficienciesallow continuous communication between machines and humans which gives room for water treatment operators to relate with AI models smoothly.This easy communication channel streamlines the decision-making procedures which enable water treatment operators to have easy and direct access to critical information and respond punctually to evolving water quality cases. The ability to process, gather and analyze data in real-time enhances water treatment facilities to take active steps which preventpossiblewaterborne diseases and ensure an uninterrupted and safe drinking water supply.

Additionally, the potential for predictive optimization and maintenance in water treatment processes is often unlocked via AI-Driven and data analytics solutions. AI algorithms and historical data can be harnessed into water treatment facilities in order to assess performance patterns, predict equipment malfunctions and proactively plan the maintenance activities. Predictive maintenance reduces downtime and alleviatesunexpected breakdowns risk and thus ensures efficient and continuous water treatment plants operation. Also, optimization algorithms which can adjust treatment parameters dynamically with reference to real-time data can be offered by AI-Driven and data analytics systems. The proficiency to adjust treatment processes to different water compositions reduces wastage, optimizes chemical dosages and enhances treatment efficiency. As a result of this, substantial cost savings with the minimization of treatment operations environmental impact can be attained by water treatment facilities. The interaction between water treatment technologies and data-driven AI also enhances intelligent process control and thus, allow autonomous treatment processes adjustments [15]. Data trends are analyzed by AI-Driven and data analytics solutions and make informed decisions on water treatment techniques in order to ensure compliance with regulatory standards and consistent water quality. Water treatment facilities can operate with improved precision with the achievement of a greater treatment performance level using data-driven decision-making and greater automation. Artificial neural network (ANN) is a sub-set of the AI data-driven tools with prevailing attributes and wide applications in water treatment facilities. Figure 2 is the structure and arrangement of a typical ANN.The ANN deep architecture allows it to learn and train hierarchical data representations via the combination of lower-level features in successive layers to learnhigher-level kinds. In most cases, ANN usually has a largernumber of hidden neurons and layers which are commonly utilized to capture sophisticated patterns in data and learn robust non-linear mappings from inputs to outputs [16].

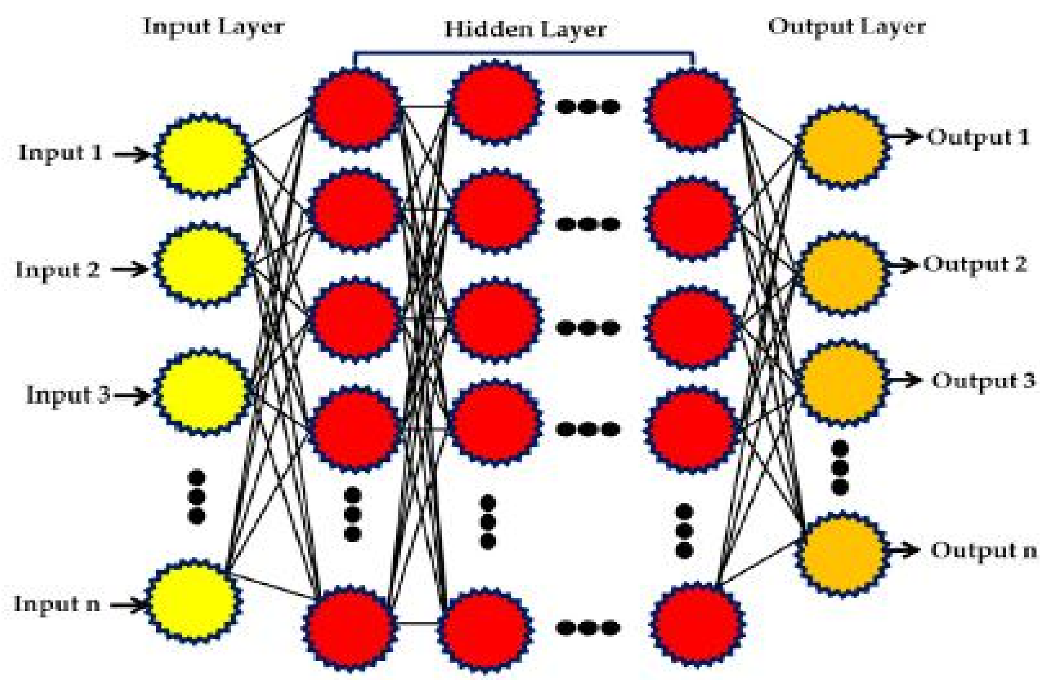

Figure 2: Structure and arrangement of a typical ANNI

2.0 SAFETY PROTOCOLS IN WATER TREATMENT FACILITIES

A water treatment facility depends on a variety of chemicals, equipment and procedures that pose serious threats to workers. Thus, it is highly imperative to equip workers with the essentialknowledge and skills to avoid accidents in the water treatment facility plant. Some major protocols have to be in place in water treatment facilities.

2.0.1 Required Safety Steps

It is vital that workers should wear appropriate PPE at all times to prevent being a victim of an unexpected accident. This prevents water treatment facility workers from exposure to dangerous biological agents, harmful chemicals and related physical hazards. It is imperative that workers should put on goggles, gloves, respirators and other PPE to serve as defense for them against the aforementioned dangers. Also, workers should strictly follow safety guidelines and procedure in handling water treatment facilities. The rules and regulations should be pasted in strategic locations within the plant for workers to be reading, understand and follow them strictly. Aside this, wearing of loose clothes or jewelries which could get trapped by machinery should be avoided when handling water treatment facility. This may cause serious injuries. Also, workers should be aware of the potential hazards that are attributed with machinery operation and chemicals handling [17-19].

Appropriate precautions should be taken to curb accidents. Its good if workers don't take everything for granted and assume things are safe. Wide varieties of machinery and chemicals are being used in water treatment plants which can inflict injuries if not properly handled. It is also essential that all chemicals should be properly labeled and stored in order to avoid accidental exposure and workers should not allow carelessness to be the cause of their demise. They should make sure that all chemicals are correctly labeled and adequately stored in appropriate locations. Workers in water treatment facility plant should strictly follow the manufacturer's rules and instructions relating to chemicals handling and storage. They should avoid mixing of chemicals without safety precautionary measures in order to avoid explosion within the plant that may lead to catastrophic consequences [18].

2.0.2 Lifting Techniques and Handling of Equipment

Additionally, workers should follow proper lifting techniques of heavy objects to prevent injury within the plant. Musculoskeletal disorders and back injuries are usually the ailments attributed with this. Proper lifting techniques should be adopted like keeping the back straight, bending at the knees and using the legs to lift heavy objects.It is well-known that debris and clutter can cause tripping hazards which could result into serious calamities. Thus, water treatment facility workers should ensure that the work areas are clutter-free and clean in order to minimize accidents risk. They should ensure that equipment is properly stored when not in use to avoid accidents. Furthermore, casual workers are not authorized to operate machinery unless they have been properly trained. Machinery operation in a water treatment plant needs authorization after specialized training. The manufacturer's instructions should be strictly adhered to on the use of the equipment only for its planned purpose. Nonetheless, workers should use caution when working near water or at specified heights. Locations and areas like this can be risky in a water treatment plant. Unnecessary risks should not be taken and appropriate safety measures should be used to prevent accidents. For instance, a harness should be worn when working at heights and barricades should be used to prevent to falling into open tanks. Lastly, any incidents or safety concerns should be reported to a supervisor instantly in order to ensure a safe working environment [20].

2.1 DATA DRIVEN TECHNIQUES IN WATER TREATMENT FACILITIES

Data-driven "intelligent" applications now cut across all areas of living. Innovative water utilities can derive some benefits from this current digital technology revolution to increase their performance. Water utilities can optimize the data and the available information in making better decisions via harnessing the distinct benefits of big data analytics and artificial intelligence algorithms while reducing costs and enhancing service delivery. Artificial intelligence (AI) is made up of "a branch of computer science handling intelligent behavior simulation in computers. Machine learning or AI is mainly used for decision-making tasks in the context of efficient water supply delivering. These include the reduction of operating costs (OPEX) and capital investment (CAPEX) optimization together with environmental and social externalities. The application of AI in water supply has numerous policy implications to increasewater utilities performanceand the service delivery quality. Governance and ethics handle the protection of financial and personal data from consumers, and financial and technical data from water utilities. Regulation handles benchmarking because unaccounted-for-water (UFW) is a major operational parameter used in evaluating water utility efficiency for the reduction of both commercial losses (like metering errors and illegal connection) and physical losses (likepipe bursts and water leaks) [12]. Figure 3 represents some main components of data driven techniques in water treatment facilities. It comprises of the network monitoring, hydraulic model, sensor data, burst detection and sensor failures together with leakage estimation. All these work together to ensure a perfect data driven water treatment system [21].

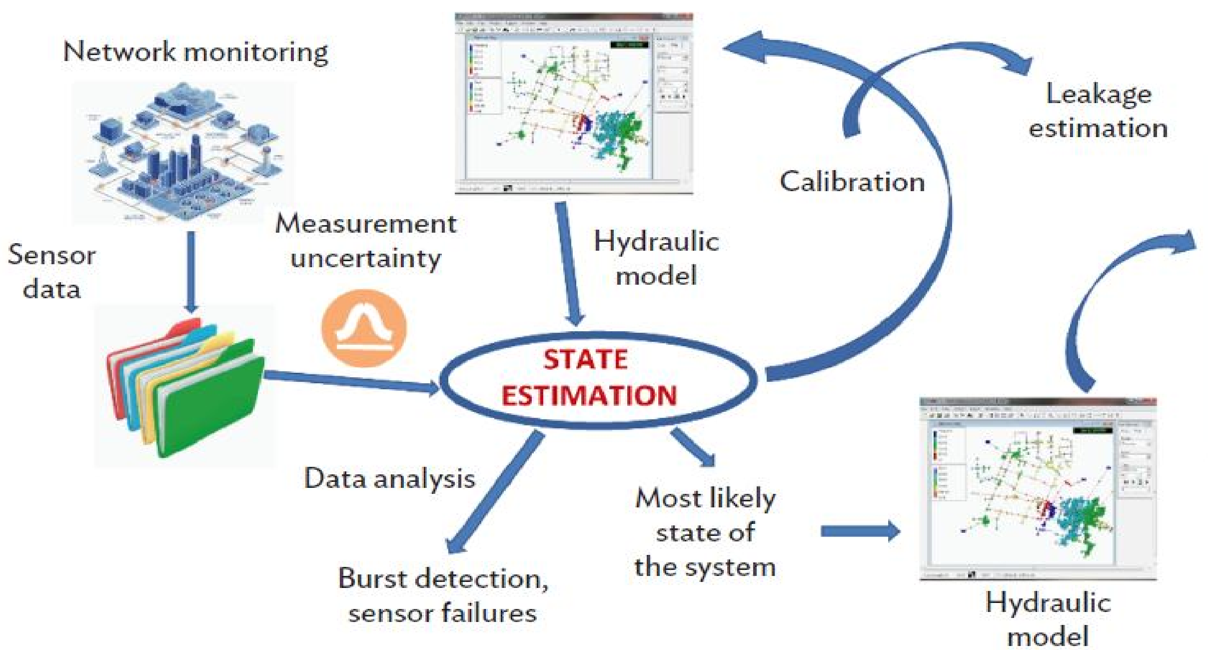

Figure 3: Some main components of data driven techniques in water treatment facilities

Technical policies handlewater associations and line ministries to update the national and water "Code of Practice," in order to ensure proper guidance on water utilities with the aid of digital transformation. Financial policies handle the requirements of the short-term CAPEX so as to finance smart water utilities using the recent financing instruments. Data-driven methods involve the use of machine learning algorithms or AI as artificial neural networks (with their different types generally called deep learning methods), as well as classification trees, support vector machines, adaptive neuro-fuzzy inference systems and so on. These methods, after being properly trained using large volume of data sets, can detect patterns and get information without using network equations. Several data-driven techniques used in the detection of pipe burst in water systems have been stated in literatures [11]. Data-driven techniques are not the natural method used in water networks hydraulic analysis (such as pipe bursts, water leaks, UFW and so on) owning by some numerous reasons. Among these are that historical databases are often too minute to be used in the detection of current abnormal events that are not previously measured and in the training of the algorithms. There are also numerous key factors influencing the hydraulic model which are complicated and not subject to the usual general governing equations thus, not applicable for physically based techniques.

Among these, cumulative human behavior, influenced by exogenous factors such as socioeconomic and climate factors, determines the illegal connections, water consumption configurations, social response to service incidences and so on. However, pipe bursts, water leaks and roughness are functions of the complicated physicochemical interactions existing between soil, pipes, water and external loads. Numerous phenomena such as pipe cracking, corrosion, joint wearing, biofilming and so on are yet to be fully understood and cannot be excellently modeled [16]. Data-driven methods can now bring further advantages for the management and operations of the water utilities when the aforementioned factors come into play. Eventually, both the data-driven and physically based techniques need to be merged together to form a “hybrid methodology”.

2.2 MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE TECHNIQUES WATER TREATMENT APPLICATIONS

Numerous water treatment applications have examined machine learning (ML) and artificial intelligent (AI) techniques. The intelligent design systems of water treatment and its reuse can be beneficial from the application of AI models together with traditional methods and IoT architecture. AI models are powerful and valuable tools for the modeling, prediction and optimization of the wastewater- and water-treatment processes. They have extensively been applied in different aspects of water treatment such as removal of heavy metals, microbial contaminants, colors, solids, organic materials, medication, pesticides and nutrients from water. The major areas in which data-driven and AI models are found useful include laboratory-scale research and process design. Process performance prediction and process parameter optimization are majorly included in process design in real-life applications. The three popular treatment techniques that are usually used in water and wastewater-treatment facilities include adsorption procedures, membrane-filtration procedures and disinfection and chlorination by-product management.Adsorption techniques are broadly identified as chemical and physical treatment options for various contaminants removal from wastewater and water treatments. Adsorption entails target molecule transfer (adsorbate) from a fluid to a solid surface (adsorbent) via an exothermic mass transfer process.However, accurate performance prediction and vital parameters calculation of the adsorption process can be tasking as a result of the complicated relations involved. ML models have been extensively applied in enhancing the adsorption process via the provision of critical predictions [17].

Membrane processes are usually adopted in the removal of contaminants that need an extreme removal level such as those that are costly or not easy to remove via other means. Popular membrane techniques include reverse osmosis, ultrafiltration, microfiltration and nanofiltration. Equally, ANNs are the principal model applied in wastewater/water-treatment applications. Disinfection is a critical technique of wastewater- and water-treatment plants that has to do with the elimination of viruses and bacteria via the utilization of chlorine-based disinfectants. Though chlorination is effective at disinfection, it poses health hazards. ML technology has been promising in mitigating and predicting the generation of disinfection by-products (DBPs) in drinking water together with controlling their levels with the aid of AI technologies. ANNs have been widely tested as ML models for DBP prediction and chlorination together with additional applications which adopt fuzzy inference systems, vector machines and genetic algorithms. In conclusion, the integration of intelligent technology with assessed AI approaches or ML models assists in the mitigation of data collection issues. Some ML models may become more accurate with more data. While ML models are beneficial in the simulation of disinfection by-product concentrations and vital parameters for membrane-filtration and adsorption operations, AI approaches have shown their efficacy in chlorination controlling [18]. Figure 4 is the applications of various ML processes in water treatment plant of various facilities.

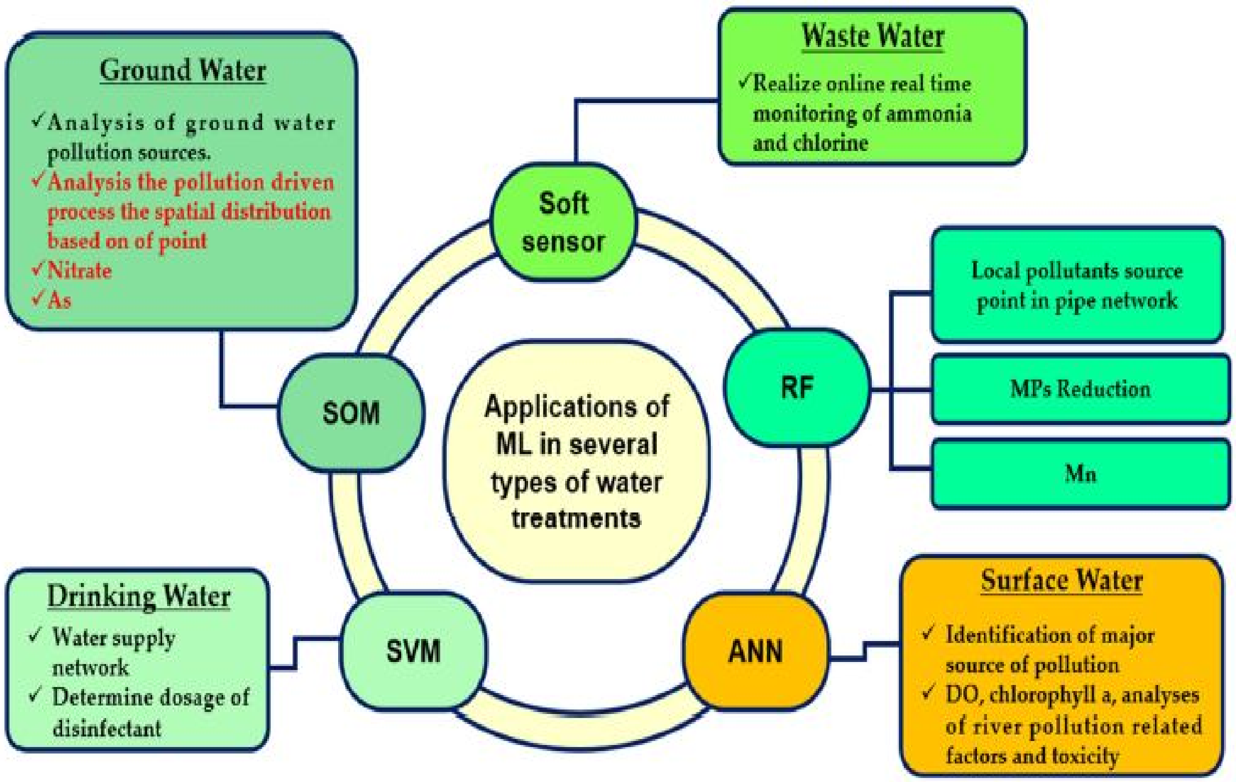

Figure 4. Applications of various ML processes in water treatment plant of various facilities

3.0 ARTIFICIAL INTELLIGENCE AS DATA ANALYTICS OPTIMIZATION TOOL FORWATER TREATMENT FACILITIES

In its most modest way, decision making utilizes optimization techniques in findingfactors combination that minimize or maximize a numerical objective function. Optimization techniques have been found applicable to water treatment facilities and water distribution networks. With the rise in cost-effective sensors and fast processors which result in inexpensive computer power, optimization strategies are now incorporated into most water utilities operation and water treatment facilities.One route is via observability analysis, or location optimization of a restricted number of sensors in order to make provision for the maximum number of information regarding the water treatment and supply system. Another is via operational analysis which involves optimized and real-time control of some parts of the water treatment facility and water distribution channel such asflow or pressure control valves available in pumping stations to ensure reduction of energy consumption or water quality monitoring after the utilization of the water treatment facility [19]. Together with the increased graphics and computer processing abilities, big data created via mobile phones, social mediaand the Internet of Things (IoT) serve as feed into AI which create a big data ecosystem. The term AI is regarded as any form algorithm form which can learn and process data, such that its performance is increased over time as it is effectively trained. Numerous statistical models such asnonhierarchical classification techniques, decision trees and Bayesian networks are currently the main support to machine learning tools which are fed by big data [22-24].

In recent time, a set of algorithms referred to as“deep learning” or “multilayered artificial neural networks” have generatedextraordinary outputs for numerous well-known applications such as voice recognition, image classification and autonomous vehicles. However, it is still vague whether deep learning will transformwater treatment facility and water network analysis in the same way it is overturningbusiness processes and customer relations. Presently, data-driven and other deep learning techniques do not appear to be a feasiblesubstitute to physically based models for network analysis. Though it has been established that AI-based algorithms for water treatment optimization, its facility management and water demand forecasting are becoming more significant than the standard ones, a “hybrid” approach which is compressed in Hydraulic Modeling 2.0 has been identified as the most preferred technique. The following functionalities are provided by AI algorithms with access to big data from the water treatment facility [25-27].

- Numerical detection of apparent and physical water losses in the water treatment facility: Spatial information on the type and volume of water losses can be provided by data-driven and AI algorithms with the help of stochastic optimization and state estimation techniques. The most likely network status is found by the AI algorithms after assigning a particular uncertainty level to already existing data. A probabilistic and continuous network calibration are performed by the AI algorithms allowing structure analysis of the errors (difference between the model predictions and the measurements) at each control point with information extraction from the error patterns. There is possibility of the distinction among various kinds of water losses (for example unauthorized consumption versus pipe leaks) which are influenced by the frequency and density of measurement in each network sector. Water leaks pinpointing or field equipment connections cannot be replaced by this numerical location of losses. However, it saves money and time from optimization of the distribution networks’ sectorization, deployment of leak detection teams and prioritization of pipe replacement programs. Additionally, the detection of abnormal measurements like equipment failure is also possible. The hydraulic model representation is improved by the probabilistic calibration via the computation of the most likely pipe roughness coefficients [28].

- Demand forecasting and consumption patterns classification: Data-driven and AI algorithms can improve and learn as many data are becoming available for forecasting of water treatment and demand at a single node or a set of nodes based on advanced statistical tools and historical data. Forecasts can be generated in real time for the next day or for the longer period measured in years to help with capacity expansion protocols. All forecasts consist of uncertainty level with reference to the volume of historical information present for the calibration of the hydraulic model. Long-term forecasting is linked to future socioeconomic and climatic situations [23].

- Optimal design of control and monitoring networks: Control and monitoring networks are the digitalized section of the physical pipes which allow water utilities transition to the digital information era. The objective information-based criteria are usually provided by the data-driven and AI algorithms in order to identify the location of a stated number of sensors within a particular network. This is purposely for the extraction of the maximum volume of information regarding the entire system with the minimum CAPEX. This indicates not only prioritizing pressure gauges installation instead of expensive flow meters but also minimizing the number of control points. This quantitative technique to network instrumentation also builds a direct link between expected operational gains and ICT investments so as to set a realistic framework for a cost–benefit evaluation which is usually missing [29].

- Incorporation of optimal configuration with network expansion design: Advanced AI optimization tools give understanding of the most effective configurations with reference to cost minimization (CAPEX together with discounted OPEX) or any other related and selected target. Dedicated data-driven and AI algorithms can specify optimal options for network expansions. They consider the uncertainty of certain design variables such as population forecast and spatial urban growth to improve a more robust technique to decision making [30].

- Energy savings: Energy savings in the water treatment facility network operations can be guided by the AI algorithms via stochastic optimization methods having two different approaches. The first involves stating the most effective operating steps which comply with the minimum service levels with a given predetermined configuration of storage tanks, pumping facilities and energy tariffs. The second involves the identification of the most cost-effective investment in a given system (increased storage capacity, pump replacement, change of energy contract and so on) for energy savings [31].

- Contingency protocols and plans definition: In an ideal way, water utilities and water treatment facilities should be prepared to handle emergencies and reduce the impacts on the customers. Additionally, abnormal cases may come up from pipe breakdown, bursts, energy blackouts, contamination events and water scarcity. The AI algorithms help in optimizing a response with respect to the level of risks. Such contingency plans should be predefined (for instance reservoirs algal blooms) or determined in real time (state the valves to be closed in order to reduce the impact of a pipe burst to the consumers) [22].

- Programming of active asset management: Numerous water utilities and water treatment facilities have a clear technique of combining replacement and maintenance to aid optimal service levels and reduce costs. Active asset management has to do with acting first, rather than responding to external random events. Algorithms and data-driven techniques specify the optimal schedules for replacing and monitoring assets with respect to the statistical explanation of their criticality, useful life and other variables [32-35].

3.1 FUTURE PROSPECTS

The use of data-driven techniques and data analytics in water treatment facilities approaches have been influential in modeling, improving and automating the required processesin wastewater treatment; monitoring of water-based agriculture together with natural systems management. The integration of data-driven techniques and data analytics is expected to lower the costs, offer computer-assisted solutions for sophisticated challenges and equally enhance water-based applications.in relation to physical/biological processes and water chemistry. Data-driven techniques and data analytics have successfully modeled,predicted, automated, and optimized significant applicationsin water-related operations and industries including wastewater- and water-treatmentfacilities, water-based agriculture and natural systems. it is recommended to conduct future studies on the improvement of water infrastructure resiliency viadata-driven techniques and data analytics and create soft sensors for water-treatment plants.

3.2 LIMITATIONS AND CHALLENGES

Despite the aforementioned future prospects of data-driven techniques and data analytics in water treatment facilities, there are some certain limitations and challenges which are needed to be adequately looked into. Issues such as availability of data and its quality have been very challenging. A substantial volume of high-quality data having high precision is required which is very rare in water-treatment facilities and management systems as a result of technological or financial constraints. Secondly, there is limited applicability of the data driven process as a result of their complex conditions usually encountered in real wastewater management and treatment systems. This implies that there may be limited suitability of the present methods for specific systems. There is need to address challenges that are related to data management with reference to its transparency and repeatability in research studies. While these limitations and shortcomings are evident, ongoing current research and development reveal the significantpotential and implications of data-driven techniques and data analytics in one of the world's most vital resources, water.

4. CONCLUSION

The most significant part of human existence and industrial operations has been linked to water which often becomes vulnerable to hazardous contaminants which are brought on through natural processes and human activity. This challenge has necessitated the use of water treatment facilities in order to make water a sustainable natural resource for the whole world. To achieve this, it is imperative to monitor and classify water quality. However, the conventional methods of water treatment have been identified with numerous challenges in terms of deficiencies in record keeping and their utilization for improved treatment processes. This has called for revolutionized water treatment facilities operation and optimization enhanced with the utilization of data-driven techniques and harnessing data analytics. This review article has critically examined harnessing of data-driven and data analytics in water treatment facilities with consideration also given to innovative safety protocols and process optimization. The basic fundamental principle of water treatment facilities were discussed alongside the application of data-driven techniques. Safety protocols in water treatment facilities were discussed. Data driven techniques in water treatment facilities and applications of machine learning and artificial intelligence techniques to water treatment were explained. Consideration was given to artificial intelligence as data analytics optimization tool for water treatment facilities. In conclusion, AI algorithms and historical data can be harnessed into water treatment facilities in order to assess performance patterns, predict equipment malfunctions and proactively plan the maintenance activities.

Disclaimer (Artificial intelligence)

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc) and text-to-image generators have been used during writing or editing of manuscripts.

REFERENCES

1. Bahramian, M., Dereli, R.K., Zhao, W., Giberti, M., Casey, E., 2022. Data to intelligence: the role of data-driven models in wastewater treatment. Expert Syst. Appl., 119453

2. Ray, S.S.; Verma, R.K.; Singh, A.; Ganesapillai, M.; Kwon, Y.-N. A holistic review on how artificial intelligence has redefined water treatment and seawater desalination processes. Desalination 2023, 546, 116221.

3. Shen, Y.; Chen, C.; Li, P.; Huang, X.; Li, Y. Application of a smart pilot electrochemical system for recycling aquaculture seawater. Aquac. Int. 2023, 1–17.

4. Martínez, R.; Vela, N.; El Aatik, A.; Murray, E.; Roche, P.; Navarro, J.M. On the use of an IoT integrated system for water quality monitoring and management in wastewater treatment plants. Water 2020, 12, 1096.

5. Karn, A.L.; Pandya, S.; Mehbodniya, A.; Arslan, F.; Sharma, D.K.; Phasinam, K.; Aftab, M.N.; Rajan, R.; Bommisetti, R.K.; Sengan, S. An integrated approach for sustainable development of wastewater treatment and management system using IoT in smart cities. Soft Comput. 2021, 27, 5159–5175.

6. Lowe, M.; Qin, R.; Mao, X. A review on machine learning, artificial intelligence, and smart technology in water treatment and monitoring. Water 2022, 14, 1384.

7. Suquet, J.; Godo-Pla, L.; Valentí, M.; Verdaguer, M.; Martin, M.J.; Poch, M.; Monclús, H. Development of an environmental decision support system for enhanced coagulation in drinking water production. Water 2020, 12, 2115.

8. Xie, Y.; Chen, Y.; Lian, Q.; Yin, H.; Peng, J.; Sheng, M.; Wang, Y. Enhancing real-time prediction of effluent water quality of wastewater treatment plant based on improved feedforward neural network coupled with optimization algorithm. Water 2022, 14, 1053.

9. Jagtap, S.; Skouteris, G.; Choudhari, V.; Rahimifard, S. Improving water efficiency in the beverage industry with the internet of things. In Implementing Data Analytics and Architectures for Next Generation Wireless Communications; IGI Global: Hershey, PN, USA, 2022; pp. 18–26.

10. Singh, N.K.; Yadav, M.; Singh, V.; Padhiyar, H.; Kumar, V.; Bhatia, S.K.; Show, P.-L. Artificial intelligence and machine learningbased monitoring and design of biological wastewater treatment systems. Bioresour. Technol. 2022, 369, 128486.

11. Altowayti,W.A.H.; Shahir, S.; Othman, N.; Eisa, T.A.E.; Yafooz,W.M.; Al-Dhaqm, A.; Soon, C.Y.; Yahya, I.B.; Che Rahim, N.A.N.b.; Abaker, M. The role of conventional methods and artificial intelligence in the wastewater treatment: A comprehensive review. Processes 2022, 10, 1832.

12. Qi, J.; Hou, Y.; Hu, J.; Ruan, W.; Xiang, Y.; Wei, X. Decontamination of methylene Blue from simulated wastewater by the mesoporous rGO/Fe/Co nanohybrids: Artificial intelligence modeling and optimization. Mater. Today Commun. 2020, 24, 100709.

13. Rodríguez-Rángel, H.; Arias, D.M.; Morales-Rosales, L.A.; Gonzalez-Huitron, V.; Valenzuela Partida, M.; García, J. Machine learning methods modeling carbohydrate-enriched cyanobacteria biomass production in wastewater treatment systems. Energies 2022, 15, 2500.

14. Guo, Z.; Du, B.; Wang, J.; Shen, Y.; Li, Q.; Feng, D.; Gao, X.; Wang, H. Data-driven prediction and control of wastewater treatment process through the combination of convolutional neural network and recurrent neural network. RSC Adv. 2020, 10, 13410–13419.

15. Lee, M.W.; Hong, S.H.; Choi, H.; Kim, J.-H.; Lee, D.S.; Park, J.M. Real-time remote monitoring of small-scaled biological wastewater treatment plants by a multivariate statistical process control and neural network-based software sensors. Process Biochem. 2008, 43, 1107–1113.

16. Park, J.; Lee, C.H.; Cho, K.H.; Hong, S.; Kim, Y.M.; Park, Y. Modeling trihalomethanes concentrations in water treatment plants using machine learning techniques. Desalinat. Water Treat. 2018, 111, 125–133.

17. Rashid, R.; Shafiq, I.; Akhter, P.; Iqbal, M.J.; Hussain, M. A state-of-the-art review on wastewater treatment techniques: The effectiveness of adsorption method. Environ. Sci. Pollut. Res. 2021, 28, 9050–9066.

18. Haghiabi, A.H.; Nasrolahi, A.H.; Parsaie, A. Water quality prediction using machine learning methods. Water Qual. Res. J. 2018, 53, 3–13.

19. Ohale P.E., Igwegbe C.A., Iwuozor K.O., Emenike, E.C., Obi, C.C., Białowiec, A. A review of the adsorption method for norfloxacin reduction from aqueous media. MethodsX, 2024,102180.

20. Krittanawong, C. The rise of artificial intelligence and the uncertain future for physicians. Eur. J. Intern. Med. 2024, 48, e13–e14.

21. Suquet, J.; Godo-Pla, L.; Valentí, M.; Verdaguer, M.; Martin, M.J.; Poch, M.; Monclús, H. Development of an environmental decision support system for enhanced coagulation in drinking water production. Water, 2020, 12, 2115.

22. Wu, Y.; Zhang, X.; Xiao, Y.; Feng, J. Attention neural network for water image classification under IoT environment. Appl. Sci. 2023, 10, 909.

23. Jagtap, S.; Skouteris, G.; Choudhari, V.; Rahimifard, S. Improving water efficiency in the beverage industry with the internet of things. In Implementing Data Analytics and Architectures for Next Generation Wireless Communications; IGI Global: Hershey, PN, USA, 2022; pp. 18–26.

24. Altowayti,W.A.H.; Shahir, S.; Othman, N.; Eisa, T.A.E.; Yafooz,W.M.; Al-Dhaqm, A.; Soon, C.Y.; Yahya, I.B.; Che Rahim, N.A.N.b.; Abaker, M. The role of conventional methods and artificial intelligence in the wastewater treatment: A comprehensive review. Processes 2022, 10, 1832.

25. Bahramian, M., Dereli, R.K., Zhao, W., Giberti, M., Casey, E.. Data to intelligence: the role of data-driven models in wastewater treatment. Expert Syst. Appl., 2024, 119453

26. Sherstinsky, A. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. Phys. D Nonlinear Phenom. 2023, 404, 132306.

27. Mahesh, B. Machine learning algorithms-a review. Int. J. Sci. Res. (IJSR) 2022, 9, 381–386.

28. Chen, K.; Chen, H.; Zhou, C.; Huang, Y.; Qi, X.; Shen, R.; Liu, F.; Zuo, M.; Zou, X.; Wang, J. Comparative analysis of surface water quality prediction performance and identification of key water parameters using different machine learning models based on big data. Water Res. 2020, 171, 115454.

29. Emenike, E.C., Adeleke, J., Iwuozor, K.O., Ogunniyi, S., Adeyanju, C.A., Amusa, V.T. Adsorption of crude oil from aqueous solution: a review. J. Water Process Eng. 2023, 50, 103330.

30. George, A.S., George, A.H., Martin, A.G.. The environmental impact of AI: a casestudy of water consumption by chat GPT. Partners Univ. Int. Innov. J. 2024, 1(2), 97–104.

31. Gursoy, D., Li, Y., Song, H. ChatGPT and the hospitality and tourism industry: anoverview of current trends and future research directions. J. Hosp. Mark. Manag. 2023, 32 (5), 579–592.

32. Moon, T.; Kim, Y.; Kim, H.; Choi, M.; Kim, C. Fuzzy rule-based inference of reasons for high effluent quality in municipalwastewater treatment plant. Korean J. Chem. Eng. 2011, 28, 817–824.

33. Lowe, M.; Qin, R.; Mao, X. A review on machine learning, artificial intelligence, and smart technology in water treatment and monitoring. Water 2022, 14, 1384.

34. Qi, J.; Hou, Y.; Hu, J.; Ruan, W.; Xiang, Y.; Wei, X. Decontamination of methylene Blue from simulated wastewater by the mesoporous rGO/Fe/Co nanohybrids: Artificial intelligence modeling and optimization. Mater. Today Commun. 2020, 24, 100709.

35. Rodríguez-Rángel, H.; Arias, D.M.; Morales-Rosales, L.A.; Gonzalez-Huitron, V.; Valenzuela Partida, M.; García, J. Machine learning methods modeling carbohydrate-enriched cyanobacteria biomass production in wastewater treatment systems. Energies 2022, 15, 2500.