How to Cite:
Puli, S., Bindu, N. ., Partha, B., Meghana, K., & Anirudh, V. (2022). Quality risk analysis for sustainable smart water supply using data perception. *International Journal of Health Sciences*, *6*(S5), 5270–5278. https://doi.org/10.53730/ijhs.v6nS5.9826

Quality risk analysis for sustainable smart water supply using data perception

Mrs. Srilatha Puli
Assistant Professor, CSE Sreyas Institute of Engineering and Technology, Telangana, India.
Email: srilatha.puli@sreyas.ac.in

Nandyala Bindu
Department of CSE, Sreyas Institute of Engineering and Technology, Telangana, India.
Email: nandyalabindu971@gmail.com

Belide Partha
Department of CSE, Sreyas Institute of Engineering and Technology, Telangana, India.
Email: parthabelide@gmail.com

Katta Meghana
Department of CSE, Sreyas Institute of Engineering and Technology, Telangana, India.
Email: meghanakatta26@gmail.com

Vaddula Anirudh
Department of CSE, Sreyas Institute of Engineering and Technology, Telangana, India.
Email: anirudh.vaddula@gmail.com

***Abstract*---**Constructing Sustainable Smart Water Supply systems are facing serious challenges all around the world with the fast expansion of modern cities. Water quality is influencing our life ubiquitously and prioritizing all the urban management. Traditional urban water quality control mostly focused on routine tests of quality indicators, which include physical, chemical and biological groups. However, the inevitable delay for biological indicators has increased the health risk and leads to accidents such as massive infections in many big cities. In this paper, we first analyse the problem, technical challenges, and research questions. Then we provide a possible solution by building a risk analysis framework for the urban water supply system. It takes indicator data we collected from industrial processes to perceive water quality changes, and further for risk detection. In order to provide

Manuscript submitted: 27 Feb 2022, Manuscript revised: 9 April 2022, Accepted for publication: 18 June 2022

explainable results, we propose an Adaptive Frequency Analysis (Adp-FA) method to resolve the data using indicators' frequency domain information for their inner relationships and individual prediction. We also investigate the scalability properties of this method from indicator, geography and time domains. For the application, we select industrial quality data sets collected from a Norwegian project in 4 different urban water supply systems, as Oslo, Bergen, Strømmen and Aalesund. We employ the proposed method to test spectrogram, prediction accuracy and time consumption, comparing with classical Artificial Neural Network and Random Forest methods. The results show our method better perform in most of the aspects. It is feasible to support industrial water quality risk early warnings and further decision support.

***Keyword*---**smart water, chemical, data perception.

Introduction

During the latest years of 21st century, two important phenomena have been emerging: urbanization and information technologies. The United Nations (UN) Department of Economic and Social Affairs (DESA) reports that for the first time ever, the majority of the world's population lives in cities, and this proportion continues to grow with projections of 68% by 2050 [1]. Urban water supply systems are the most critical infrastructure all over the world. A Smart Water Supply system that integrates sensors, controllers, cloud computing and data technologies, are essential for the development of sustainable smart cities in the future. It is aiming to provide safe, stable and sufficient water for the increasing requirements in many expanding cities. However, the urban water quality is facing serious challenges from industrial, agriculture and social pollution. To emphasize the importance of water safety in urban supply is nowadays a truism. In 2015, the United Nations Development Programme published the Sustainable Development Goals (SDGs), including Clean Water and Sanitation as Goal 6 [2]. The dwindling supplies of safe drinking water is a major problem impacting every continent, around 2.1 billion people [3]. The concerns of the modern society regarding this issue are reflected in numerous legislative initiatives in this field, such as the European Union Water Framework Directive [4], United States Clean Water Act [5]. The prevalent water supply process can be divided into 3 sections, including water source management, treatment, and distribution. Traditional water quality control is taken after water treatment. But the current water sources are mainly groundwater and surface water. They are significantly prone to chemical and microbial contamination. The quality control after the water treatment apparently delays the risk detection and reduces the response time to take preventive measures. In Norway, the new national standard for water quality in the source area is in progress [6] [7]. Water quality refers to physical, chemical, and biological characteristics as indicators. Among the water quality indicators, biological indicators have a more direct impact over people's health. Most of the national standards are made on biological indicator levels. Typical indicators include coliform, Escherichia coli (Ecoli), intestinal enterococci (Int), clostridium perfringens (ClPerf), etc. Further treatment actions are made according to the test results [8]. Coliform itself is not usually causing serious illness, but their presence is a signal to indicate other active pathogenic organisms presentation. Some special types of Ecoli are the reason for water poisoning. Int is more dangerous to cause urinary tract infections, bacterial endocarditis, diverticulitis, and meningitis. The tests of biological indicators are primarily based on the bacterial culture in the 2 laboratory. This process can take up to 24-48 hours. Compare to the effectual time on the human body, the danger is much higher than other indicators. In Norway, The giardia outbreak in Bergen 2004 affected more than 2500 people including young children due to the bacteria test delay results. Therefore, we have a severe requirement for early risk detection in smart water supply systems. There have been some trial work for water quality control based on data. In 2018, Hounslow [9] interpreted multiple water quality indicators. In 2015, Yagur-Kroll et al [10] showed a group of general bacterial sensor cells for water quality monitoring. There is some research work to use data for water quality prediction. Holger et al[11] designed an Artificial neural network to predict salinity level for an Australian river named Murray. Based on the data collected at Astane station in Sefidrood River, Iran, Orouji and his colleagues designed a series of models as ANFIS, GA and Shuffled FLA to predict water quality chemical indicators (sodium, potassium, magnesium, etc) in [12] [13] [14]. Chang et al[15] proposed a systematic analysis framework to predict NH3-H levels for DahanRiver in Taiwan, China. However, their work is generally on individual quality indicator and ignored the inner relationship between them. Today the advanced ubiquitous sensing technologies cut across many areas of modern research, industry and daily life [16]. They offer the ability to detect, transmit and measure more environmental indicators. A sustainable smart water supply system adopts various sensors in order to manage resources and monitor water quality efficiently. In this process, data becomes an important tool to improve our understanding of existing systems. By observing data itself, through the appropriate methods, we can perceive the changes in our water supply system. In practice, we applied many different sensors in the water source areas, including multiple sensors for pH, temperature, conductivity,etc. The massive data collected by those low-cost sensors plus the recent data analysis technologies, help us greatly improve the water quality control process. At present, zettabytes of data are collected by these numerous sensors [17] [18]. At the same time, stronger data analysis tools have been developed. Water quality indicators are typical spatiotemporal variables. The analysis can be divided into correlation analysis and numerical prediction analysis. Early works with correlation analysis include Hardoon et al [19] used Kernel Correlation Analysis method for web page images and associated texts. For multiple variables,Principal component analysis(PCA) is often the first choice. Jolliffe et al [20] reviewed classical PCA and newly developed methods such as Robust PCA, Adaptive PCA, etc. Luo et al [21] applied tensor model in correlation analysis for gait recognition. But they did not consider the correlations in the time domain. As for spatiotemporal data analysis, most of the recent work is facing very huge datasets. For example, Gudmundsson et al [22] surveyed the player's trajectories in team-sports 3 with respect to behaviour and prediction. Lecun et al [23]proposed the pioneer concept for Deep Learning to deal with spatiotemporal data. Liu et al [24]analysed 3D human actions with modern LSTM method. Laptev et al [25] detects anomalies in the industrial platform data. However, their work has to rely on large training sets, which we cannot provide currently in water supply systems. In addition, the explanation with those methods cannot support the requirements for industrial use. In this paper, we introduce our preliminary experience in Norway. First, we analyse the problem, challenges and research questions. Second, based on water quality data collected from water supply systems, we propose a framework for water quality analysis with data perception. Third, we provide an adaptive frequency analysis method for risk detection and prediction. This method is scalable in multiple domains, including water quality indicators, geography and time. Furthermore, by application, we select industrial quality data sets collected from a national project in 4 different Norwegian city water supply systems, as Oslo, Bergen, Strømmen and Alesund. We show our preliminary ˚ findings of the frequency property relationship between water quality indicators, as well as risk detection, prediction and evaluation analysis. The results are compared also with classical Artifical Neural Network and Random Forest in their prediction accuracy and time consumption. In addition, scalability in time domain is also analyzed. There are several visible motivations for this research. First, it takes the advantage of the modern data analysis technologies to solve a water quality control problem in future Sustainable Smart Water Supply systems, especially in transferring the knowledge across different indicator, geography and time domains. Second, it copes with the practical water source monitoring process, applies thedata directly collected from the industrial process. This avoids questions such as laboratory data reliability and industrial applicability. This is also valuable to the current water supply in urban infrastructure systems. Third, it builds the connection between easily accessible physical and chemical indicators with biological indicators that are critical to water quality risk. Fourth, this work provides the support for further reasoning of decision-making process and analysis over the pollution from industrial and residential activities in the corresponding water source areas.

Literature Survey

Urbanization and climate change impacts on surface water quality Authors: S. Franco, V. Gaetano, and T. Gianni Abstract: Climate change and urbanization are key factors affecting thefuture of water quality in urbanized catchments. The work reported in this paper is an evaluation of the combined and relative impact of climate change and urbanization on the water quality of receiving water bodies in the context of a highly urbanized watershed served by a combined sewer system (CSS) in northern Italy. The impact is determined by an integrated modelling study involving two years of field campaigns. The results obtained from the case study show that impervious urban surfaces and rainfall intensity are significant predictors of combined sewer overflows (CSOs) and consequently of the water quality of the receiving water body. Scenarios for the year 2100 demonstrate that climate change combined with increasing urbanization is likely to lead to severe worsening of river water quality due to a doubling of the total phosphorus load from CSOs compared to the current load. Reduction in imperviousness was found to be a suitable strategy to adapt to the scenarios by limiting the construction of new impervious areas and decreasing the existing areas by only 15%. This information can be further utilized to develop future designs, which in turn should make these systems more resilient to future changes in climate and urbanization. Sustainable development goals: A need for relevant indicators Authors: T. Hak, S. Janoskova, and B. Moldan Abstract: At the UN in New York the Open Working Group created by the UN General Assembly proposed a set of global Sustainable Development Goals (SDGs) which comprises 17 goals and 169 targets. Further to that, a preliminary set of 330 indicators was introduced in March 2015. Some SDGs build on preceding Millennium Development Goals while others incorporate new ideas. A critical review has revealed that indicators of varied quality (in terms of the fulfilment certain criteria)have been proposed to assess sustainable development. Despite the fact that there is plenty of theoretical work on quality standards for indicators, in practice users cannot often be sure how adequately the indicators measure the monitored phenomena LITERATURE SURVEY Urbanization and climate change impacts on surface water quality Authors: S. Franco, V. Gaetano, and T. Gianni Abstract: Climate change and urbanization are key factors affecting the future of water quality in urbanized catchments. The work reported in this paper is an evaluation of the combined and relative impact of climate change and urbanization on the water quality of receiving water bodies in the context of a highly urbanized watershed served by a combined sewer system (CSS) in northern Italy. The impact is determined by an integrated modelling study involving two years of field campaigns. The results obtained from the case study show that impervious urban surfaces and rainfall intensity are significant predictors of combined sewer overflows (CSOs) and consequently of the water quality of the receiving water body. Scenarios for the year 2100 demonstrate that climate change combined with increasing urbanization is likely to lead to severe worsening of river water quality due to a doubling of the total phosphorus load from CSOs compared to the current load. Reduction in imperviousness was found to be a suitable strategy to adapt to the scenarios by limiting the construction of new impervious areas and decreasing the existing areas by only 15%. This information can be further utilized to develop future designs, which in turn should make these systems more resilient to future changes in climate and urbanization. Sustainable development goals: A need for relevant indicators Authors: T. Hak, S. Janoskova, and B.

Moldan Abstract: At the UN in New York the Open Working Group created by the UN General Assembly proposed a set of global Sustainable Development Goals (SDGs) which comprises 17 goals and 169 targets. Further to that, a preliminary set of 330 indicators was introduced in March 2015. Some SDGs build on preceding Millennium Development Goals while others incorporate new ideas. A critical review has revealed that indicators of varied quality (in terms of the fulfilment certain criteria)have been proposed to assess sustainable development. Despite the fact that there is plenty of theoretical work on quality standards for indicators, in practice users cannot often be sure how adequately the indicators measure the monitored phenomena.

Existing System and Proposed System

Existing ANN and random forest will not have above dataset processing steps so its error rate will be high compare to propose adaptive frequency analysis algorithm. All existing algorithms will not filter dataset multiple times to extract important features which helps in getting better prediction accuracy and reduce error rate. Disadvantages • High error rate. • Data preprocessing steps are not found in Existing System.

In this project as extension we have added Adaptive Frequency Analysis Algorithm. This method to resolve the data using indicators frequency domain information for their inner relationships and individual prediction. We also investigate the scalability properties of this method from indicator, geography and time domains. Advantages • Reduce error rate. • Better prediction accuracy. • Increasing performance and popularity

Fig 1 Displaying risk prediction accuracy and RMSE

In above screen we trained all algorithms on same dataset and from above output we can see Adapative Frequency Analysis has got more accuracy and less error rate compare to all other algorithms and below is then RMSE comparison graph for all algorithms.

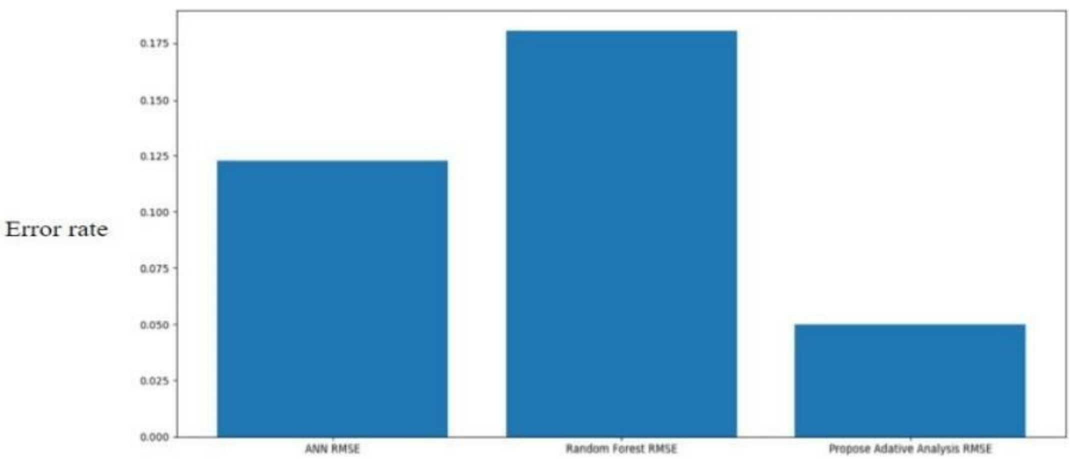

Fig 2: RMSE Comparison graph

In above graph we can see Adaptive Frequency Analysis got less error compare to other algorithms and in machine learning model with less error rate and more accuracy can be consider as best prediction model.

Fig 3: Predicted water quality risk

Conclusion

Water quality is a very critical issue in modern urban life all around the world, especially for Smart Water Supply system development. Traditional monitoring and risk control methods are difficult to detect bacteria broadcast on time and provide efficient decision support. In this paper, we propose an approach for water quality risk early warning using data perception. With the application among four different cities in Norway, we have proved the feasibility, accuracy, and efficiency of our approach. The preliminary results evaluated by domain experts are very promising. This work is beneficial in generally three aspects: It provides an early warning mechanism from the water source areas using cost-less data analysis techniques. This prolongs the preventive measures response time, and support more decision options in the latter steps of water supply. This approach integrates indicator, geography and time domains. It provides a new frequency domain analysis perspective to find the relationship between different indicators and their predictions. At the same time, it embraces scalability for these three domains. This work is applied to real industrial water supply systems from 4 different Norwegian cities. Future Scope Furthermore, we have to decluster the results and predict accurate bacteria indicators, both in tendency and values. These values can map to different risk modes according to practical water source management standards in different countries and regions. Future decision support in water treatment plants can adjust to both prediction and risk mode. Also, in practice, the models need to be evolved with both domain knowledge data set growing.

References

[1] S. Franco, V. Gaetano, and T. Gianni, “Urbanization and climate change impacts on surface water quality: Enhancing the resilience by reducing impervious surfaces,” Water Research, vol.144, pp. 491–502, 2018.

[2] T. Hak, S. Janouskova, and B. Moldan, “Sustainable development ´ goals: A need for relevant indicators,” Ecological Indicators, vol. 60, pp. 565–573, 2016.

[3] World Health Organization (WHO), Guidelines for drinking-water quality: recommendations. World Health Organization, 2004.

[4] E. Weinthal, Y. Parag, A. Vengosh, A. Muti, and W. Kloppmann, “The eu drinking water directive: the boron standard and scientific uncertainty,” European Environment, vol. 15, no. 1, pp. 1–12, 2005.

[5] R. W. Adler, J. C. Landman, and D. M. Cameron, The clean water act 20 years later. IslandPress, 1993.

[6] V. Novotny, Water quality: prevention, identification and management of diffuse pollution.Van Nostrand-Reinhold Publishers, 1994.

[7] A. Hounslow, Water quality data: analysis and interpretation. CRC press, 2018.

[8] S. Yagur-Kroll, E. Schreuder, C. J. Ingham, R. Heideman, R. Rosen, and S. Belkin, “A miniature porous aluminum oxide-based flowcell for online water quality monitoring using bacterial sensor cells,” Biosensors and Bioelectronics, vol. 64, pp. 625–632, 2015.

[9] H. R. Maier and G. C. Dandy, “The use of artificial neural networks for the prediction of water quality parameters,” Water Resources Research, vol. 32, no. 4, pp. 1013–1022, 1996.

[10] H. Orouji, O. Bozorg Haddad, E. Fallah-Mehdipour, and M. Marino, “Modeling of water quality parameters using data- ˜ driven models,” Journal of Environmental Engineering, vol. 139, no. 7, pp. 947–957, 2013.

[11] D. R. Hardoon, S. Szedmak, and J. Shawe-Taylor, “Canonical correlation analysis: An overview with application to learning methods,” Neural computation, vol. 16, no. 12, pp. 2639–2664, 2004 .

[12] I.T. Jolliffe and J. Cadima, “Principal component analysis: a review and recent developments,” Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, vol. 374, no. 2065, p. 20150202, 2016.

[13] Srilatha Puli, a machine learning model for air quality prediction for smart cities, design engineering || issn: 0011-9342 | year 2021 - issue: 9 | pages: 18090 – 18104.

[14] Srilatha Puli, n-grams assisted youtube spamcommentdetection, ymer || issn : 0044-0477, april 2022.

[15] Srilatha Puli, analysis of brand popularity using big data and twitter, ymer|| issn : 0044- 0477, april 2022.

[16] Kumar, S. (2022). A quest for sustainium (sustainability Premium): review of sustainable bonds. Academy of Accounting and Financial Studies Journal, Vol. 26, no.2, pp. 1-18

[17] Allugunti V.R (2022). A machine learning model for skin disease classification using convolution neural network. International Journal of Computing, Programming and Database Management 3(1), 141-147

[18] Allugunti V.R (2022). Breast cancer detection based on thermographic images

using machine learning and deep learning algorithms. International Journal of Engineering in Computer Science 4(1), 49-56

[19] Srilatha Puli, cyber threat detection based on artificial neural networks using event profiles, the international journal of analytical and experimental modal analysis, issn no:0886- 9367.

[20] Suryasa, I. W., Rodríguez-Gámez, M., & Koldoris, T. (2022). Post-pandemic health and its sustainability: Educational situation. International Journal of Health Sciences, 6(1), i-v. https://doi.org/10.53730/ijhs.v6n1.5949

[21] Srilatha Puli, face mask monitoring system, the international journal of analytical and experimental modal analysis, issn no:0886-9367.

[22] Srilatha Puli, iot based smart door lock surveillance system using security sensors, advanced science letters e-issn:1936-7317.

[23] N. Swapna Suhasini, Srilatha Puli, big data analytics for malware detection in a virtualized framework, journal of critical reviews, issn:2394-5125 vol.7, issue 14, july – 2020.

[24] Dr. Shaik Abdul Nabi, Srilatha Puli, framework for social distance monitoring with computer vision and yolov3, 2nd international conference on iot based control networks and intelligent systems (icicnis 2021), ssrn-id3883372.