International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

QUALITY RISK ANALYSIS FOR SUSTAINABLE SMART WATER SUPPLY USING DATA PERCEPTION

GurramDhanalakshmi[1], T.V.Gopala Krishna[2]

[1]*M.Tech student , Department of COMPUTER SCIENCE and ENGINEERING, Loyola Institute of Technology and Management, Dhulipalla (Village), Guntur (Dist), Sattenapalle, Andhra Pradesh 522412*

[2]*ASSOCIATE PROFESSOR, Department of COMPUTER SCIENCE and ENGINEERING, Loyola Institute of Technology and Management, Dhulipalla (Village), Guntur (Dist), Sattenapalle, Andhra Pradesh 522412*

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 analyze 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 explainable results, we propose an Adaptive Frequency Analysis (AdpFA) 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, Strommen 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.

Keywords:ANN, FA, RISK, QUALITY.

1. INTRODUCTION:

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. 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. 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 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 has been some trial work for water quality control based on data. In 2018, Hounslow interpreted multiple water quality indicators. In 2015, Yagur-Kroll et al. 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. 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.). Chang et al. proposed a systematic analysis framework to predict NH3-H levels for Dahan River 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. 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 appropriatemethods, 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.

2. LITERATURE SURVEY

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. • In propose paper author has used Norwegian country water supply dataset but he did not publish that dataset on internet so we don't have that dataset but we found Indian state water supply quality dataset.

DISADVANTAGES

In order to evaluate the risk from water quality change and analyze the mechanism behind the data resources, we are facing several challenges: 1. Data Sparsity: the pool of available data is often very large. In practice, for water quality indicator samples, the overlaps between two conditions (such as the same time, same location) are often very small or none. This is based on two main reasons. First, the operators who take the samples do not follow the standard procedure (incomplete indicator collections, and data loss). Second, data standard has been changed over last years (indicators have been added or removed). These make the data set sparse.

2. Data Synchronization: current sensing technologies can support real-time data collection over most of the physical and chemical indicators for water quality. However, for biological indicators, which are the key factors for health, the tests usually take much longer time, from several hours to several days. This makes the data set difficult to synchronize.

3. Risk Modeling: the final objective of drinking water quality control is to improve health. Some specific biological indicators as bacteria can cause significant disease outbreaks, such as Ecoli.

When they broadcast in the drinking water distribution system, the consequences can be irreversible. The relationship between those biological indicators and drinking water risk needs a new model.

3. METHODOLOGY

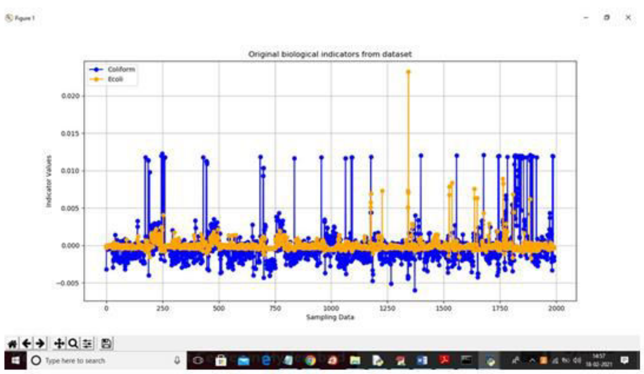
In this project as extension, we have added CNN (convolution neural network) and LSTM (long shortterm memory) and compare RMSE (root mean square error) with existing algorithms such as Random Forest, ANN and Adaptive Frequency. All existing algorithms will not filter dataset multiple times to extract important features which helps in getting better prediction accuracy and reduce error rate. CNN and LSTM are the two most preferable deep learning algorithms which filter dataset multiple times to extract important features from dataset and then train a prediction model and all irrelevant features will be removed out by using DROPOUT functions and dataset will be filtered using function called DENSE which filtered dataset by using specified number of neurons. More data processing organizations are switching towards CNN and LSTM classification or prediction model due to its increasing performance and popularity. In CNN and LSTM, we will define number of input and output layers and each layer will take number of data filtration as input. In below code screen you can read red color comments to understand CNN implementation.

After the data is prepared, we need to find the key factors from multiple dimensions of indicators by primary correlations analysis, probability distribution and generate training and testing data sets. The eventual aim of this work is to predict water quality risk. In order to find the risk model, we have investigated with researchers from water quality control. Here the risk evaluation model is further divided into three parts. Cycle detection is to find the hidden cycle for indicator changes in the time domain. Peak value calculation is used to track and evaluate the levels of multiple biological bacteria outbreaks. Parameter correction is based on training set adaptation. 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.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

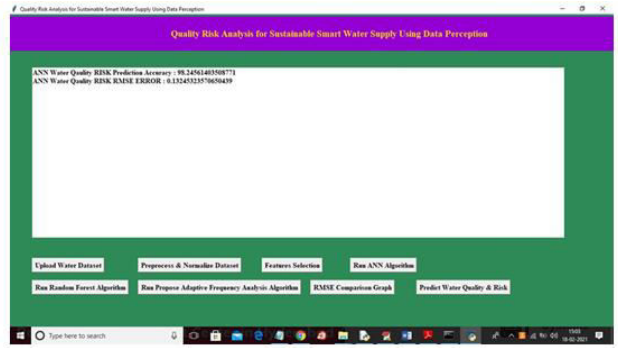
HospitalDatabase

Inabovescreenbeforeapplyingfeatureselectiondatasetcontaining12attributesandafterapplyingfeatureselection algorithm attributes reduceto9and then will getbelow graph



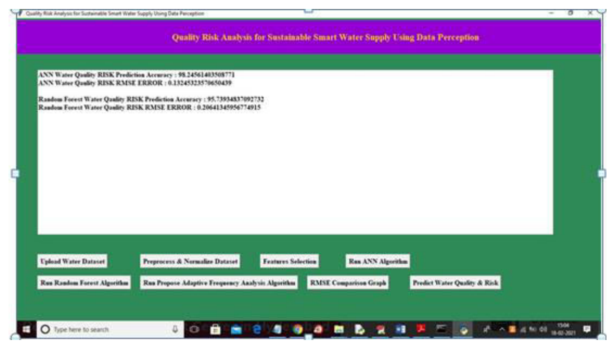
AISearch String

In above graph we can see COLIFORM and ECOLI bacteria present in water dataset where blue colourrepresents presence of COLIFORM and orange colour represents COLI bacteria present in dataset. Now datasetisreadyandnowclickon‘Run ANNAlgorithm’buttontotrainANNon abovedataset andcalculateRMSE



AISearch Result

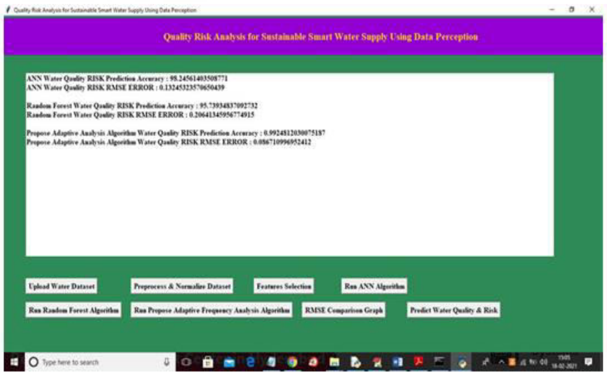
Inabove screenANN accuracy is98% and itsRMSEis0.13% andnow clickon ‘Run RandomForestAlgorithm’buttontotraindatasetwith randomforestand get belowresult



PatientLogin

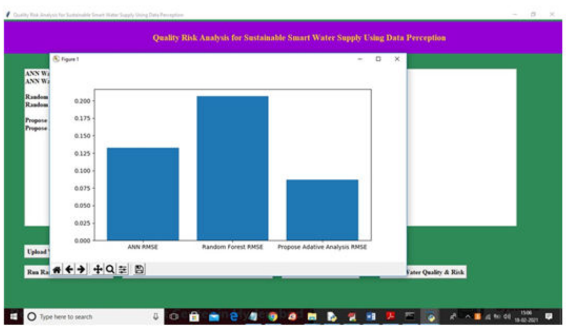
InabovescreenRandomForestAccuracyis95.73% anditserrorrateis0.20%andnowclickon‘RunProposeAdaptiveFrequencyAnalysisAlgorithm’buttontotrainalgorithm withdatasetandget belowresult

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

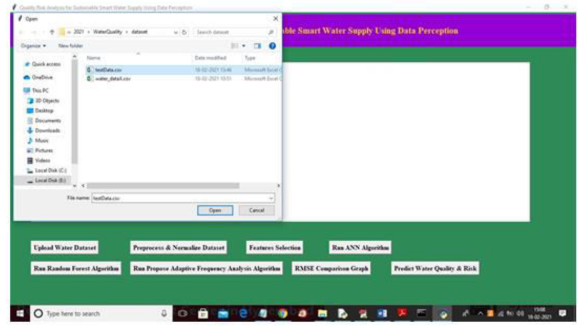


PatientDetails

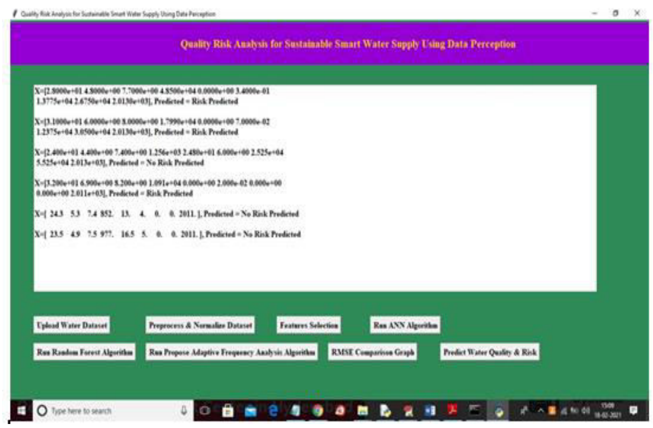
In above screen propose algorithm accuracy is 0.99 * 100 = 99% and its error rate is 0.086% which is lesser thanabovetwo algorithmsandnowclickon‘RMSE ComparisonGraph’ buttonto getbelowgraph



Inabovegraphx-axisrepresentsalgorithmnameandy-axisrepresentsRMSEerrorrateandfromall3algorithms propose Adaptive algorithm got less error rate. Now click on ‘Predict Water Quality & Risk’ buttontouploadtest water dataand thenapplication will predictwhetherthatwaterdata isRISKYornot



In above screen selecting and uploading ‘testData.csv’ file and then click on ‘Open’ button to get belowpredictionresult



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.

REFERANCES

[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. Janou ´ skov ˇ a, 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. Island Press, 1993.

[6] D. Berge, “Overvakingavfarrisvannet med tilløpfra 1958-2010,” ˚ 2011.

[7] I. W. Andersen, “EUs rammedirektiv for vann– miljøkvalitetsnormer for vannmiljøetimøte med norskrett,” Kart og Plan, vol. 73, no. 5, pp. 355–366, 2013.

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

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

[10] 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.

[11] 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.

[12] 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.