ISSN 2394-739X

IJRSET MAY Volume 10 Issue 5

International Journal for Research in Science Engineering & Technology (IJRSET)

https://www.doi.org/10.5281/zenodo.8366824

WIRELESS WATER QUALITY MONITORING AND DETERIORATION PREDICTION SYSTEM

[1] Dr. G. Jawaherlalnehru, [2] Dr. R. SanthoshKumar, [3] Dr. B. Rajalingam,
[4] Dr. M. Vadivukarassi, [5] S. Bavankumar
[1,2,3,4] Associate Professor, St. Martins Engineering College, Secunderabad, Telangana.
[5]Assistant Professor, St. Martins Engineering College, Secunderabad, Telangana.

ABSTRACT: Water is an essential resource in day-to-day life. Pollution and urbanization have led to higher susceptibility of source water to contamination. There is a pressing need to develop a water quality monitoring system to preserve the quality of source water and ultimately safeguard human health. This proposes a low cost, wireless water quality monitoring system, wherein the quality of water stored in overhead tanks is continuously monitored. The quality of water is measured by parameters that are critical quality indicators. The data encompassing these parameters are stored in a Cloud database (in realtime) along with its timestamp. The quality of water is ascertained based on the comparison of the monitored data to standard well-established thresholds. The data, annotated with its timestamp is treated as a time-series. A univariate non-seasonal Auto Regressive Integrated Moving Average (ARIMA) model is employed to forecast individual water quality parameters. The results of forecasting are used to predict water quality deterioration. The model used is found to have mean square errors of 0.001 for pH, 0.076 for temperature and 0.001 for turbidity between the actual and forecasted values.

Keywords: [Auto Regressive Integrated Moving Average (ARIMA)]

1. INTRODUCTION

Water used for commercial or domestic purposes is commonly stored in overhead tanks. The water being stored can become a breeding ground for various pathogens or harmful microorganisms. Contact with rainwater alters acidity, rendering this water unfit for consumption and other purposes. In the long run, harmful chemicals may be deposited on the walls of the tank. Exposure to open air can lead to contamination by particulate matter. Sedimentation of these particles can alter certain chemical properties of the water. Formation of rust resulting from improper maintenance of water collection pipes severely degrades the quality of water. Microbiological quality of water characterizes illhealth. Infectious diseases like dysentery, giardiasis, cholera, typhoid, guinea worm, hepatitis, and schistosomiasis are spread through water contamination.

These diseases may result from inadequate sanitation or poor hygiene. It is important to note that all factors regarding quality and availability of drinking-water is important and is implicated in public health. As a precursor to tackling these problems, a real-time system is proposed which continuously monitors and reports the quality of the water. The data monitored by the system truly captures the water quality. A useful analysis of this data can be used to predict the deterioration of water quality. This can be achieved by employing time-series forecasting. Auto Regressive Integrated Moving Average (ARIMA) is a forecasting statistical technique that is used to analyze a time series based on history. A non-seasonal model is well-suited as it must be insensitive to any local shortlived trends within the time series. These trends do not contribute to the overall quality of water in the future.

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. 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.

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 coli form, 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.

It 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 have 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, 3 magnesium, etc) in. 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 appropriate methods, we can perceive the changes in our water supply system.

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 the data 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.

2. LITERATURE SURVEY

1) "Urbanization and climate change impacts on surface water quality: Enhancing the resilience by reducing impervious surfaces," Climate change and urbanization are key factors affecting the future of water quality in urbanized catchments. The work reported in this project 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 these 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.

2)"Sustainable development ´ goals: A need for relevant indicators," 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. Therefore we stress the need to operationalise the Sustainable Development Goals' targets and evaluate the indicators' 5 relevance, the characteristic of utmost importance among the indicators' quality traits. The current format of the proposed SDGs and their targets has laid a policy framework; however, without thorough expert and scientific follow up on their operationalisation the indicators may be ambiguous. Therefore we argue for the foundation of a conceptual framework for selecting appropriate indicators for targets from existing sets or formulating new ones. Experts should focus on the "indicator-indicated fact" relation to ensure the indicators' relevance in order for clear, unambiguous messages to be conveyed to users (decision- and policy-makers and also the lay public). Finally we offer some recommendations for indicators providers in order to contribute to the tremendous amount of conceptual work needed to lay a strong foundation for the development of the final indicators framework.

3)"The eu drinking water directive: the boron standard and scientific uncertainty," In 1998 the European Union (EU) revised its Drinking Water Directive, which is responsible for regulating the quality of water in the EU intended for human consumption. Specifically, the EU added a new standard for the element boron in drinking water (1 mg/l). Yet, because of scientific uncertainty concerning the causes and magnitude of the boron problem in Europe during the regulatory standard-setting process, we find that full compliance with the new drinking water standard for boron has been hampered. Prior to the standard's enactment, it was unclear whether boron was derived from natural or anthropogenic sources. A new geochemical study reveals that a significant part of the boron contamination is derived from natural sources. Countries such as Italy and Cyprus with high natural boron concentrations in their drinking water are, thus, finding that compliance with the new EU boron regulation is more difficult and expensive than originally anticipated.

4) "The use of artificial neural networks for the prediction of water quality parameters," This project presents the use of artificial neural networks (ANNs) as a viable means of forecasting water quality parameters. A review of ANNs is given, and a case study is presented in which ANN methods are used to forecast salinity in the River Murray at 6 Murray Bridge (South Australia) 14 days in advance. It is estimated that high salinity levels in the Murray cause $US 22 million damage per year to water users in Adelaide. Previous studies have shown that the average salinity of the water supplied to Adelaide could be reduced by about 10% if pumping from the Murray were to be scheduled in an optimal manner. This requires forecasts of salinity several weeks in advance. The results obtained were most promising. The average absolute percentage errors of the independent 14-day forecasts for four different years of data varied from 5.3% to 7.0%. The average absolute percentage error obtained as part of a real-time forecasting simulation for 1991 was 6.5%

5**) "Modelling of water quality parameters using data-driven models**," Water has a considerable role in all aspects of human life. Thus, evaluation of water characteristics in general and water quality in particular are necessary to enhance the health of humans and ecosystems. The writers considered six combinations of data sets, including the previously noted water quality parameters and river discharge in the previous and previous-current months, as input data. Implementation of the ANFIS and GP models in this project illustrates the flexibility of GP in time series modeling relative to ANFIS, especially in the testing data set. Accordingly, the writers calculated the coefficient of variation of root mean squared error as the error criterion in different ANFIS and GP models (for assigning achievement probability to an appropriate solution) for each quality parameter. The average of the previously noted values for the six combinations of data sets improved (decreased) 80.51 and 80.89%, respectively, in the training and testing data sets with GP relative to ANFIS. These results indicate that the writers' proposed modelling, based on GP, is an effective tool for determining water quality parameters. (C) 2013 American Society of Civil Engineers.

6)"Modeling ´ water-quality parameters using genetic algorithm–least 7 squares support vector regression and genetic programming," The modeling and monitoring of water-quality parameters is necessary because of the ever increasing use of water resources and contamination caused by sewage disposal. This study employs two data-driven methods for modeling water-quality parameters. The methods are the least-squares support vector regression (LSSVR) and genetic programming (GP). Model inputs to the LSSVR algorithm and GP were determined using principal component analysis (PCA). The genetic algorithm (GA) was combined with LSSVR to produce the GA-LSSVR algorithm with which to achieve improved accuracy in modeling water-quality parameters. The GA-LSSVR algorithm and the GP method were modeling Na+Na+, K+K+, Mg2+Mg2+, SO2−4SO42−, Cl−Cl-, pH, electric conductivity (EC), and total dissolved solids (TDS) in the Sefidrood River, Iran. The results indicate that the GA-LSSVR algorithm has better accuracy for modeling waterquality parameters than GP judged by the coefficient of determination (R2R2) and the NS criterion. The NS static established, however, that the GA-LSSVR and GP methods have the capacity to model water-quality parameters accurately.

7) "Integration of shuffled frog leaping algorithm and support vector regression for prediction of water quality parameters," Quality of surface water is a serious factor affecting human health and ecological systems. Accurate prediction of water quality parameters plays an important role in the management of rivers. This study presents a novel tool for estimation of quality parameters by coupling SVR and shuffled frog leaping algorithm (SFLA) . Results of SFLA-SVR compared with genetic programming (GP) as a capable method in water quality prediction. Using SFLA-SVR, average of RMSE for training and testing of six combinations of data sets for all of the water quality parameters improved 57.4 % relative to GP. These results indicate that the new proposed SFLA-SVR tool is more efficient and powerful than GP for determining water quality parameters.

8)"Modeling water quality in an urban river using hydrological factors– data driven approaches," Contrasting seasonal variations occur in river flow and water quality as a result of short duration, severe intensity storms and typhoons in Taiwan. Sudden changes in river flow caused by impending extreme events may impose serious degradation on river water quality and fateful impacts on ecosystems. Water quality is measured in a monthly/quarterly scale, and therefore an estimation of water quality in a daily scale would be of good help for timely river pollution management. This study proposes a systematic analysis scheme (SAS) to assess the spatio-temporal interrelation of water quality in an urban river and construct water quality estimation models using two static and one dynamic artificial neural networks (ANNs) coupled with the Gamma test (GT) based on water quality, hydrological and economic data.

9) "Principal component analysis: a review and recent developments," Large datasets are increasingly common and are often difficult to interpret. Principal component analysis (PCA) is a technique for reducing the dimensionality of such datasets, increasing interpretability but at the same time minimizing information loss. It does so by creating new uncorrelated variables that successively maximize variance. Finding such new variables, the principal components, reduces to solving an eigenvalue/eigenvector problem, and the new variables are defined by the dataset at hand, not a priori, hence making PCA an adaptive data analysis technique. It is adaptive in another sense too, since variants of the technique have been developed that are tailored to various different data types and structures. This article will begin by introducing the basic ideas of PCA, discussing what it can and cannot do. It will then describe some variants of PCA and their application.

10) "Tensor canonical correlation analysis for multi-view dimension reduction," Canonical correlation analysis (CCA) has proven an effective tool for two-view dimension reduction due to its profound theoretical foundation and success in practical applications. In respect of multi-view learning, however, it is limited by its capability of only handling data represented by two-view features, while in many real-world applications, the number of views is frequently many more. As a consequence, the high order correlation information contained in the different views is explored and thus a more reliable common subspace shared by all features can be obtained. In addition, a non-linear extension of TCCA is presented. Experiments on various challenge tasks, including large scale biometric structure prediction, internet advertisement classification, and web image annotation, demonstrate the effectiveness of the proposed method.

3. PROPOSED SYSTEM

An approach to data resolution known as Adaptive Frequency Analysis (Adp-FA) that makes use of the frequency domain information of indicators for the resolution of their inner relationships and individual prediction. We also look into the scalability of this method across the indicator, geographical, and temporal dimensions. Industrial quality data sets collected as part of a Norwegian project in four urban water supply systems (Oslo, Bergen, Strmmen, and Alesund) are used in this application. We put the proposed method through its paces by utilising it to perform spectrogram, prediction accuracy, and time consumption tests against traditional Artificial Neural Network and Random Forest approaches. The outcomes suggest that our approach outperforms alternatives in most cases. It is possible to provide early warnings for risks associated with water quality in industrial settings and continue to provide decision support in light of these warnings.

4. MODULES

In proposed algorithm we have added multiple features such as

1) Physical data. Drinking water has to verify physical attributes in water quality for the whole supply process.
2) Chemical data. Chemical indicators are the traditional representation of water quality. They provide information on what is impacting on the system as well.
3) Biological data. Biological indicators are direct measures of the health of the fauna and flora in the water supply.
4) Environmental data. Environment data can be a leading impact factor for water quality in some places.
5) Data Preprocessing: In this module missing and irrelevant data will be removed from dataset as data obtained from sensor contains huge size of data with noise values so by preprocessing we can remove such noise data.
6) Normalization: using this module we can normalize data between 0 and 1 to allow machine learning model to make better prediction
7) Clustering: Using this module we will cluster all dataset and this cluster will separate risk water quality data into one cluster and clean quality data into other cluster.
8) Synchronize: collect only recent data for evaluation

ISSN 2394-739X

Implementation and Testing:

Implementation is one of the most important tasks in project is the phase in which one has to be cautions because all the efforts undertaken during the project will be very interactive. Implementation is the most crucial stage in achieving successful system and giving the users confidence that the new system is workable and effective. Each program is tested individually at the time of development using the sample data and has verified that these programs link together in the way specified in the program specification. The computer system and its environment are tested to the satisfaction of the user.

Implementation

The implementation phase is less creative than system design. It is primarily concerned with user training, and file conversion. The system may be requiring extensive usertraining. The initial parameters of the system should be modifies as a result of aprogramming. A simple operating procedure is provided so that the user can understand the different functions clearly and quickly. The different reports can be obtained either on the inkjet or dot matrix printer, which is available at the disposal of the user. The proposed system is very easy to implement. In general implementation is used to mean the process of converting a new or revised system design into an operational one.

Testing

Testing is the process where the test data is prepared and is used for testing the modules individually and later the validation given for the fields. Then the system testingtakes place which makes sure that all components of the system property functions as a unit. The test data should be chosen such that it passed through all possible condition. Actually testing is the state of implementation which aimed at ensuring that the system works accurately and efficiently before the actual operation commence. The following is the description of the testing strategies, which were carried out during the testing period.

Quality Risk Analysis for Sustainable Smart Water Supply Using Data Perception In this project we are describing an algorithm called Adp-FA (Adaptive Frequency Analysis) to predict commercial water quality and its risk as water quality is a major concern for almost all countries due to increasing population and government wants to supply quality and sufficient water to its citizens. In this project we are analysing water quality sensor data obtained from 4 different states (OSLO, Bergen, Strommen and Alesund) of Norwegian country and simultaneously using scalable properties such as INDICATOR, Geography (locations) and Time.

In this we are building machine learning model by using water supply dataset and then comparing propose Adp-FA algorithm RMSE (root mean square error) with ANN and Random Forest. Propose algorithm giving less RMSE error rate compare to ANN and Random Forest.

After processing dataset using above points then we will split dataset into train and test and then decluster data to predict risk value. In dataset if 100ML water contains any value of In this project, we had 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 and below screen shots showing values from dataset and this dataset saved inside 'dataset' folder

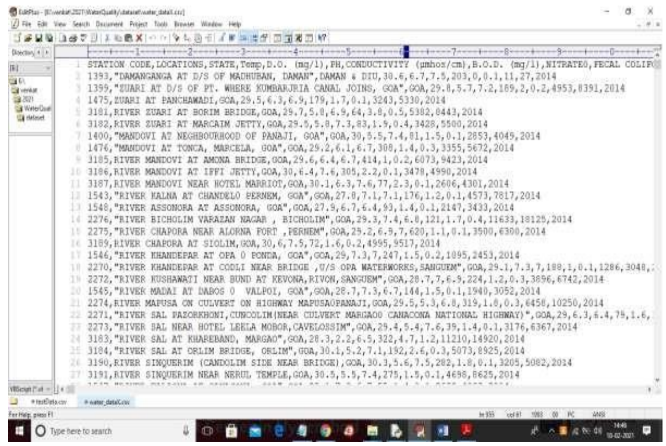

Figure. 1 : dataset folder

Figure. 2 : Main screen

Figure.3: ANN and RSME accuracy

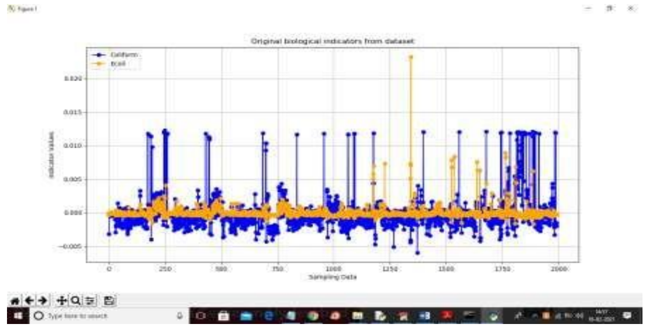

Figure.4: Graph

In above screen in square brackets we can see water test data and after square bracket we can see predicted result as ‘Risk Predicted’ or ‘No Risk Predicted’ in water. This prediction is coming by evaluating water data such as ECOLI presence, temperature value, PH and many other test values.

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 proposed 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 ingenerally 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 workis applied to real industrial water supply systems from 4 different Norwegian cities

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. 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. This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/. This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSUSC.2019.2929953, IEEE Transactions on Sustainable Computing IEEE TRANSACTIONS ON SUSTAINABLE COMPUTING. 12

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

[6]. D. Berge, “Overvaking av farrisvannet med tilløp fra 1958-2010,” ˚ 2011.

[7]. I. W. Andersen, “EUs rammedirektiv for vann–miljøkvalitetsnormer for vannmiljøeti møte med norsk rett,” 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.