



Monitoring and forecasting water consumption and detecting leakage using an IoT system

Jyoti Gautam , Amlan Chakrabarti , Shruti Agarwal, Anushka Singh, Shweta Gupta and Jatin Singh

ABSTRACT

Water is an important resource for life and its existence and, unfortunately, large quantities of water are being wasted on a daily basis. Monitoring the consumption of water can control water usage, and smart technologies can play a useful role. In this paper, a smart system based on Internet of Things (IoT) has been proposed to monitor the water consumption in an urban housing complex. An ultrasonic sensor, together with Arduino, continuously monitors the water level of water tanks on rooftops and sends these data to a server through a Wi-Fi module. Using the data collected from the IoT system, the daily and weekly average water requirement of households can be calculated. Support vector machines (SVM) are used to forecast water consumption. The observed readings are divided into training and testing datasets. Water consumption is predicted for each day for a user. Error is recorded as the difference between the actual consumption and the predicted value, and it decreases as the number of days increase. An algorithm to monitor leakage of water in the tanks has also been proposed. A web interface allows the user to visualize the water usage, monitor their consumption, and detect any leakage and leakage rate in the system.

Key words | IoT, leakage detection, machine learning, smart water management, SVM, water consumption

Jyoti Gautam  (corresponding author)
CSE Department, JSS Academy of Technical Education,
NOIDA, Uttar Pradesh 201301,
India
E-mail: jyotjssaten@gmail.com

Amlan Chakrabarti 
AKCSIT, University of Calcutta,
Kolkata 700106,
India

Shruti Agarwal
Indiamart Intermesh Ltd, Assotech Business
Cresterra,
Noida 201305,
India

Anushka Singh
Tata Consultancy Services,
Haryana,
India

Shweta Gupta
Robert Bosch Engineering and Business Solutions,
chil sez it park, Coimbatore,
India

Jatin Singh
Tata Consultancy Services,
Noida, Uttar Pradesh 201309,
India

INTRODUCTION

Due to rapid urbanization and population increase, the stress on freshwater resources is ever increasing. This will have catastrophic impacts on food security and poverty (Sayre & Taraz 2019). Proper techniques should therefore be adopted for water conservation, and technology can play an important role. Water should be used optimally, i.e. just for use and no wastage. Water losses, over-use, and quality are some of the issues in urban water systems. Information and communication technology can play a significant role through the development of smart water grids that network and automate monitoring and control devices (Mutchek & Williams 2014). With the help of smart technologies, not only can the water consumption

be continuously monitored but the user can also be given control of their consumption. Real-time modeling and control of water distribution in a particular urban/rural region is mainly guided by water consumption data in various households or a group of households, and our methodology of estimating the water usage can play a useful role in this activity. Large amounts of water is also wasted through water leakage in tanks and pipes. Detecting these leakages and notifying the user can be a step toward monitoring and control of wastage of water. Water consumption of users was monitored using electromagnetic meters at an hourly time step and water losses were assessed with respect to the real case (Alvisi *et al.* 2019). Drivers, development, and

global deployment of intelligent water metering have been reviewed (Boyle *et al.* 2013). Smart metering helps water utilities to rapidly identify and take appropriate action on significant volumes of post-meter leakage occurring in cities (Britton *et al.* 2013). A framework has been presented for the classification of residential water demand modeling studies (Cominola *et al.* 2015). Implications of smart water metering have been discussed for enhanced water distribution infrastructure planning and management and can be found in Gurung *et al.* (2014) and Luciani *et al.* (2019). Urban water demand management planning, policy, and practice have also been discussed (Willis *et al.* 2011).

Ishido & Takahashi (2014) proposed a new algorithm that uses a technique for real-time leak detection in a water distribution network using real-time pressure measurements only. A proposal for design and development of a low-cost system for real-time monitoring of the water quality in Internet of Things (IoT) can be found in Vijayakumar & Ramya (2015). A completely data-driven, fully adaptive, self-learning algorithm for water demand forecasting in the short-term and also with hourly periodicity is proposed in Candelieri *et al.* (2015). Shahanas & Sivakumar (2016) discussed in detail the design of IoT systems and related analytics in the context of smart city initiatives in India. Machine learning techniques and smart city management aspects, such as smart water management that includes water demand forecasting, water quality monitoring, and also anomaly detection, have been detailed in Vijai & Sivakumar (2016). Research suggests that smart water management can be divided into three parts: (1) forecasting demand, (2) water quality monitoring, and (3) anomaly detection. A proposal for the design of a water tank monitoring system based on mobile devices has been presented by Gama-Moreno *et al.* (2016) called Interface for Monitoring Water Tanks (IRMA). Maevsky *et al.* (2017) studied system analysis of IoT and proposed a hierarchical architecture of smart systems.

Peña-Guzmán *et al.* (2016) presented a model for forecasting water demand in residential, commercial, and industrial zones in Bogotá, Colombia, using least-squares-support vector machines (LS-SVM) and they explored the effectiveness of this model over long time scales. Jiang *et al.* (2017) studied a particle swarm optimization

combined with SVM (PSO-SVM) model for daily water demand prediction. A proposal for using soft computing methods for forecasting water demands can be found in Ghalekhondabi *et al.* (2017). Gautam *et al.* (2018) proposed an IoT-based real-time monitoring of water levels in tanks using machine learning and an Android app. Analytics had been performed on Kaggle datasets for 19 households in Venice, Italy.

The Bureau of Waterworks in the Tokyo Metropolitan Government (2018) observed when practicing the prevention of leakage in Tokyo that the water pipes embedded underground are constantly subject to the danger of leakage, and when leakage does occur, these pipes constitute additional risks, including poor water flow, sagging roads, inundation, and so on. Thirteen cities, including Tokyo, Seoul, Los Angeles, and New York, agreed to make efforts to promote measures against leakage. A proposal for leakage detection and an estimation algorithm for loss reduction in water piping networks is presented in Adedeji *et al.* (2017). Water loss through leaking pipes is a major challenge to the operational service of water utilities. There has been significant financial loss and environmental pollution caused by leaking pipes. The algorithm helps detect critical segments or pipes in the network that are experiencing a higher leakage outflow. A summary of common leak detection technologies can be found in El-Zahab & Zayed (2019).

By focusing on the above research contributions, an IoT-based smart water management (SWM) has been proposed that involves an ultraviolet (UV) sensor and Arduino to detect water level in tanks. SWM includes data collection and integration using an ultrasonic sensor, data distribution using Arduino microcontroller, data analytics, forecasting, and decision support using machine learning techniques. The data from the UV sensor are sent to a server by the Wi-Fi device in the Arduino board, which can be monitored by users via a web interface. The data set used here has been collected from water tanks supplying water to 16 households in JSS Academy of Technical Education (JSSATE) NOIDA (<https://jssaten.ac.in/academics/CSE/index.php>), staff quarters. The next step after data collection is to apply proper analysis and optimization techniques to the data using machine learning algorithms to monitor the collected data. An SVM algorithm has been used to forecast the water demand.

An algorithm to check water leakage has also been proposed. For leakage to be detected, the basic assumption in the algorithm is that the user inputs whether the water is in use or not; if it is in use then the leakage cannot be predicted. Leakage has been detected when the water is not being used and during a period when there is no water supply. The leakage rate can also be calculated.

Key findings are as follows:

1. Development of an IoT-based smart system for monitoring water usage in an urban environment.
2. Proposal of a machine learning-based technique.
3. Proposal for water leakage detection based on user input and water usage data.
4. Development of a visualization tool for users to estimate, predict, and detect leakage.

METHODS

The phases included data acquisition, communication, storage, pre-processing, analyzing, and web interfaces carried out by the following:

1. Collecting water level (consumption) data from water tanks installed on rooftops at JSSATE NOIDA staff quarters.
2. Performing pre-processing steps on the data.
3. Analyzing the data and applying SVM for forecasting water consumption, leakage detection and leakage rate.
4. Designing a web interface for the users for visualization purposes.

This research was undertaken on the rooftops at JSSATE NOIDA staff quarters in Uttar Pradesh, India. There were two sets of tanks and both sets contained two tanks that were interconnected. Set 1 had two tanks containing reserve osmosis water used for drinking and cooking purposes; set 2 had two tanks containing water for other household activities. There were four tanks in total. The setup (Figure 1(c)) was installed on the tanks, which were connected to 16 households comprising of a total of 24 people. The system architecture is detailed in the next section. The initial, and most essential, part of our project is data acquisition, which is explained in detail later. Figure 1(a) shows the circuit diagram of Arduino with the ultrasonic sensor used for data

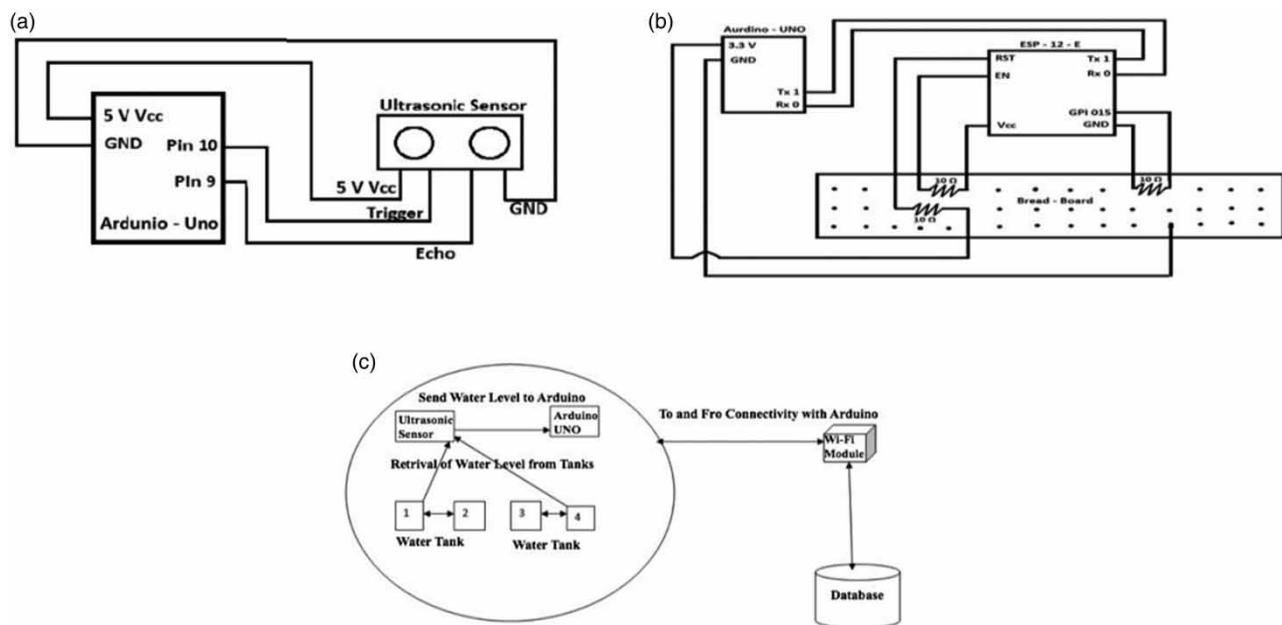


Figure 1 | Overall system architecture: (a) interfacing of Arduino with ultrasonic sensor; (b) Pin diagram of Espressif Modules (ESP) (Wi-Fi microchip) module with Arduino; and (c) data communication of Arduino with the database server.

acquisition; Figure 1(b) shows the ESP module with Arduino used for data communication; and Figure 1(c) shows the server where the data are stored. The pre-processing of data, i.e. intervals of time, the data are taken, and handling of missing values, is described later, followed by details of the analytics part, i.e., the features taken and applying the SVM and leakage detection algorithms. We then discuss the various user interfaces, starting from the login page, OTP verification page, the interface to upload a CSV file for datasets, list of attributes, the interface for total water consumption, total number of users and average consumption per user in a day, and the user interface for predicted value of water required on the next day per user on the basis of previously observed water usage. The purpose of the entire methodology is to enable users to forecast their water usage and also to inform about leakage detection through a well-suited web interface.

System architecture

The following list details the hardware components provided, along with the technical specifications for the ultrasonic sensor, data acquisition architecture, and the Wi-Fi.

Sensor and data acquisition hardware

- 1 × Breadboard
- 1 × Arduino Uno R3
- 1 × ULTRASONIC Sensor (HC-SR04)
- 1 × LCD 16 × 2

- ESP module
- Connecting wires

Technical specifications for ultrasonic sensor

Power supply +5 V DC

Quiescent current <2 mA

Working current 15 mA

Effectual angle <15°

Ranging distance 2–400 cm (1 inch–13 ft)

Data acquisition, communication, storage

The data used for the project were collected from a complex of 16 households. The height of the water tank and the requirement of the household plays an important role in determining the working of the whole process of starting and stopping the motor. The water tanks installed were Syntax 1,000 L tanks, with a 109.98-cm radius and a 122.43-cm height. There was no inflow of water in the tanks at the time of monitoring. The tank system is shown in Figure 2.

Working of ultrasonic sensor

The Arduino kit, including the UV sensor, was installed on the top of one of the two tanks to monitor the water level every 10 s. This ultrasonic sensor sends UV rays to measure the distance of the water level from the top of the tank using sonar. The sensor sends ultrasonic waves to the



Figure 2 | The water tank system used for data collection.

water, which are reflected by the water surface and returned to the sensor. The sensor uses this time of wave propagation to calculate the distance between the sensor and the water level. The time taken by the pulse is actually to and fro for travel of the ultrasonic signals, while only half of this is needed. Therefore, the time is $\text{time}/2$.

Distance calculation:

Distance = speed \times time/2

Speed of sound at sea level = 343 m/s or 34,300 cm/s

Thus the distance measured = $17,150 \times \text{time}$ (cm) (Boyle et al. 2013).

Using that distance, the water level and hence the volume of water consumed can be calculated because the height and radius of the tank are fixed.

Working of Arduino board

The Arduino is programmed by sending a set of instructions to the microcontroller on the board. The water level, which is monitored by the Arduino kit, is continuously sent to the server using the Wi-Fi module installed on the kit. The transfer of data in the system is shown in Figure 1(c).

Working of Wi-Fi module

The Wi-Fi module can connect to the Internet via a hotspot using its service set identifier (SSID) and password. It has been programmed to implement logic statements as per requirements of the project. The ultrasonic sensor reads the distance of the water surface and returns it to the module. The module, when connected to the Internet, uploads this value to the database. The water level monitored for each tank is collected by the ultrasonic sensor and the data are simultaneously transferred to the server via ESP12-e.

Calculations of current water level

In the data set, the water level detected (W) is the distance between the ultrasonic sensor and the water surface in the tank (i.e. the height of the empty space in the tank). The water height (w) is the actual height of water in the tank (i.e. the height up to which water is filled in the tank), which is calculated by subtracting the water level

detected from the total height of the water tank (cm) (Equation (1)):

$$\begin{aligned} \text{Water height (cm)} &= \text{total height (cm)} \\ &- \text{water level detected (cm)} \end{aligned} \quad (1)$$

Data description and pre-processing

Time step

The time step denotes the time interval for recording the water consumption values, and is dependent on which model is used. The model can be hourly, daily, or monthly. In our case, the time step is 10 s.

Handling missing data

In order to complete the missing values in our observed data, the mean values were taken from the observed data after observing that the consumption values did not vary greatly at the same temperature. The data were recorded for an average of 4–5 h per day over 21 days, with a total of 102 h of readings being noted. The readings were taken at the same time on different days.

Data analytics

Feature space and response variables

We considered the following parameters in defining the feature space for the SVM algorithm: (a) volume of water consumed (cm^3), (b) number of users, (c) temperature ($^{\circ}\text{C}$), and (d) precipitation (mm).

The response variables for SVM consist of: (a) average water consumption per user per day and (b) prediction of water consumption per user per day.

The algorithm for leakage detection considered the following parameters:

- (a) Input variables: water level initial, water level final and in use.
- (b) Response variables: leakage detected (yes/No) and leakage rate (cm^3/s).

SVM-based classification

SVM was used for analysis purposes because it is accurate and is the preferred method for data sets of small sizes (Ray 2017). SVM is also less prone to over-fitting than other methods and it facilitates compact models for classification. Kernel function, radial basis function (RBF), is used because the data are not linearly separable (Ng 2020) and it is the most commonly used kernel in SVM. Finally, the optimal hyper-plane is found, which creates maximum empty space at two sides of the coordinate. The independent variables used for training are shown in Table 1.

The model parameters used in designing the proposed SVM classifier are described next.

If a set independent variable is $\vec{x}_i = (x_{i1}, x_{i2}, \dots, x_{in})$ and water demand is q_i , the SVM model equation can be written as Equation (2):

$$\vec{q} = \frac{1}{2} \sum_{ij}^l \alpha_i k(\vec{x}_i, \vec{x}_j) + \beta \quad (2)$$

where q is the series of water demand prediction values and α_i is the undetermined parameters in Equation (3).

$$\alpha_i = C\xi_i \quad (3)$$

where C is penalty factor and ξ_i is slack variable,

$$k(\mathbf{x}_i, \mathbf{x}_j) = e^{-|\mathbf{x}_i - \mathbf{x}_j|/2\sigma^2} \quad (4)$$

Table 1 | Input parameters

Variables id	Variables name
X_i	Set of independent input variable
C	Penalty Factor
σ	Extension constant of RBF
α	Extra undetermined parameter
k	Kernel function
β	Offset value for SVM
V_i	Volume of water consumed (cu cm.)
T_i	Temperature (degree C)
P_i	Precipitation (mm)
N_i	Number of users
q_i	Series of water demand prediction values (from 8th day onwards)

where k is the kernel function RBF and σ is extension constant of RBF (Equation (4)).

Steps for the training algorithm

Step 1: The data set is divided into two sets: training data and testing data set. As we have data for around 25 days we take the initial 7 days' data, i.e. from 3 February 2019 to 10 February 2019, as training data.

Step 2: From the next day, i.e. 11 February 2019, we predict the total consumption values using the training data set by applying the SVM algorithm.

Step 3: The predicted value for the 8th day is compared to the actual water consumption value for that day. The error value is calculated by finding the difference between the actual and the predicted value.

Step 4: For predicting the consumption value for the next day, i.e. for the 9th day, we take the 8th days' value in the training data set. This is recursively repeated and we predict the value for n th day using the previous $n - 1$ day's value in the training data set.

Step 5: Every time we increase the data set for training data, the error value shows a significant reduction, i.e. we approach a more accurate prediction.

Steps to implement SVM

We use the following steps to predict the water consumption on the next day, i.e. using 7 days' data for learning and then increasing the number of days. The error is calculated as the difference between the predicted and the actual values (Figure 8).

We use the following steps to (predict the water consumption on the next day, i.e., using 7 days' data, for learning and increasing the number of days for learning. Calculate the error).

Step 1: Input independent input variable x (temperature, precipitation) and water demand as output y series (predicted water demand value).

Step 2: Normalize the x and y series.

Step 3: Randomly create C (penalty function) and σ (extension constant) and input the SVM model. The PSO model is used to optimize C and σ . The values are iteratively

optimized to improve solution using Equation (3) for convenience

Step 4: Use the Cross-Validation 5 (CV-5) model to train this model and calculate error E . The data set is divided into five parts using Xu & Goodacre's (2018) CV-5 model. Here one part of the data set is taken as the test data and the rest is the training data set. Then a model is fitted and error E is calculated each time (Equation (5)).

$$E = \sum_{i=1}^V \frac{1}{2} \sum_{t=1}^T (y_t - y'_t)^2 \quad (5)$$

Step 5: If E is up to the criterion, go to Step 6 or else go to Step 4.

Step 6: Output the parameters α (undetermined parameter used in the prediction value equation) and β (offset value) of SVM.

Step 7: Input new independent x_t (the next group of test data).

Step 8: Predict the new water demand in y' using Equations (2) and (4).

To implement SVM, Java Machine Learning Library (Java-ML) is used. Java-ML is a collection of machine learning algorithms with a common interface for algorithms of the same type. Java-ML contains algorithms for data pre-processing, feature selection, classification, and clustering. In addition, it features several Weka bridges to access Weka's algorithms directly through the Java-ML application programming interface.

The following libraries were used in the implementation (<http://java-ml.sourceforge.net>):

- net.sf.javaml.clustering: These are clustering algorithms such as k-means, self-organizing maps, spatial clustering, Cobweb, AQBC, and others.
- net.sf.javaml.tools: These are utility methods on data set, instance manipulation, serialization, Weka application programming interface, and so on.
- net.sf.javaml.utils: These are utility methods for algorithms, for example, statistics, math methods, contingency tables, and others.
- net.sf.javaml.classification: These are classification algorithms, including naive Bayes.

- random forest, bagging, self-organizing maps, k-nearest neighbors, and so on.

Minimizing forecasting deviation

Measuring forecast errors is crucial for the selection of an accurate and reliable forecasting model.

The basic step consists of comparing forecasts with observations, possibly using a wide subset of the available data to learn/build the forecasting model and the remaining subset of data to validate it.

The most widely adopted error measures is MAPE (mean absolute percentage error) (Equation (6)).

$$MAPE = \frac{100}{N} \sum_{i=1}^N \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right| \quad (6)$$

Y , time-series of observed water demand (at any forecast periodicity);

Y_t , water demand observed at the time t ;

\bar{Y} , average of the water demand observed;

\hat{Y} , time-series of forecasted water demand (at any forecast periodicity);

\hat{Y}_t , water demand forecasted at the time t ; and

N , time-series length.

Algorithm for leakage detection

The leakage detection system is another major feature of this project. Water leakage is a major cause of unwanted water loss, which needs to be avoided. It can be monitored and controlled only by detecting the leakages in the first place. At the basic level, the proposed leakage detection algorithm illustrated in Figure 3 can detect leakage in households. The important condition is that the water supply should not be on and the users should not be using water when the leak is being detected. Initially, the initial water level (A) is noted. After t time (which can be in seconds or minutes), the present water level (B) is noted. If $A = B$, there is no leakage; otherwise a leakage is detected and the rate of leakage is calculated. It sends a message notification to the user in case any leakage is detected in the tank and they can take actions accordingly.

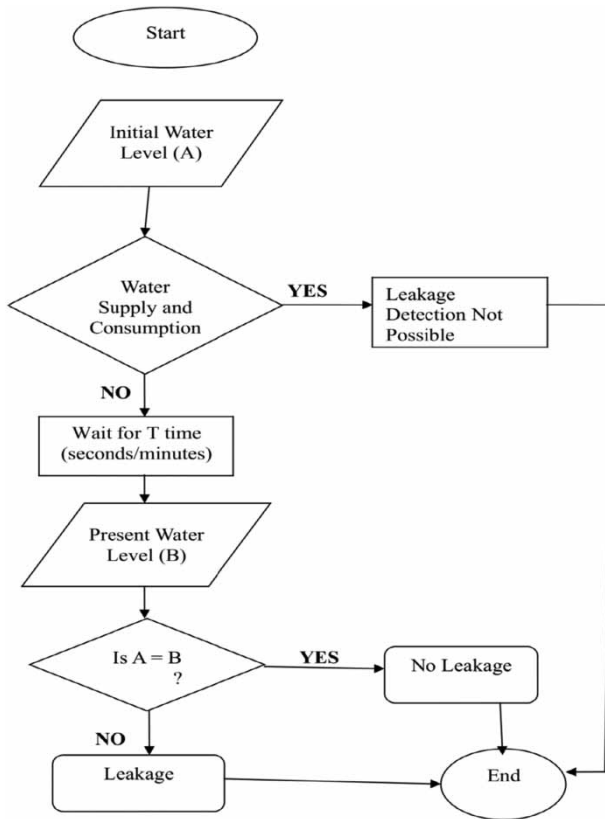


Figure 3 | Leakage detection algorithm.

The rate of leakage r is calculated using Equation (7):

$$r = (v_2 - v_1)/(t_2 - t_1) \tag{7}$$

where v_2 is final volume, v_1 is initial volume, t_2 is final time, and t_1 is initial time.

User interface design

In order to make the system more interactive for users, the project included a user interface for our smart system. The users can enter the required data, including number of current users and whether the water is in use or not. Further processing and results are shown based on the user input.

The interfaces are shown in Figures 4–6.

Figure 4 shows the interface for total water consumption, total number of users, and average consumption for a user in a day.

Figure 5 shows the user interface for the predicted value of water required on the next day per user based on the previously observed water usage. This can provide the user with a forecast of their water usage.

Figure 6 shows a snapshot of the web interface for the leakage detection system, showing whether a leakage has

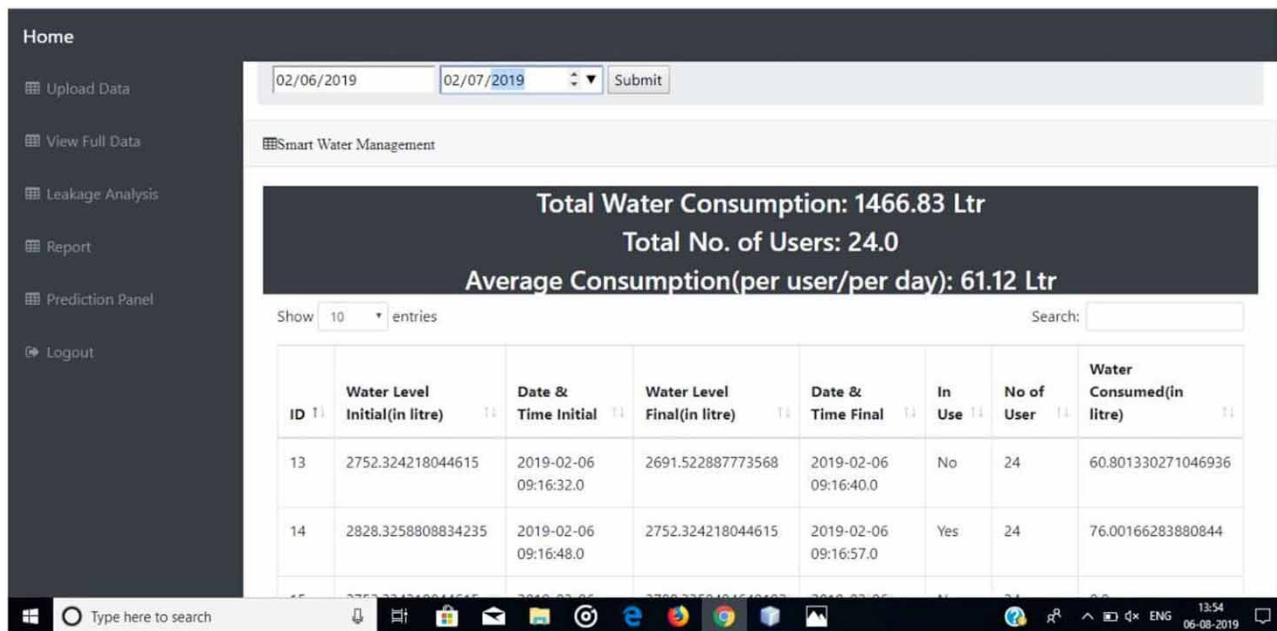


Figure 4 | Interface for average water consumption by a user in a day.

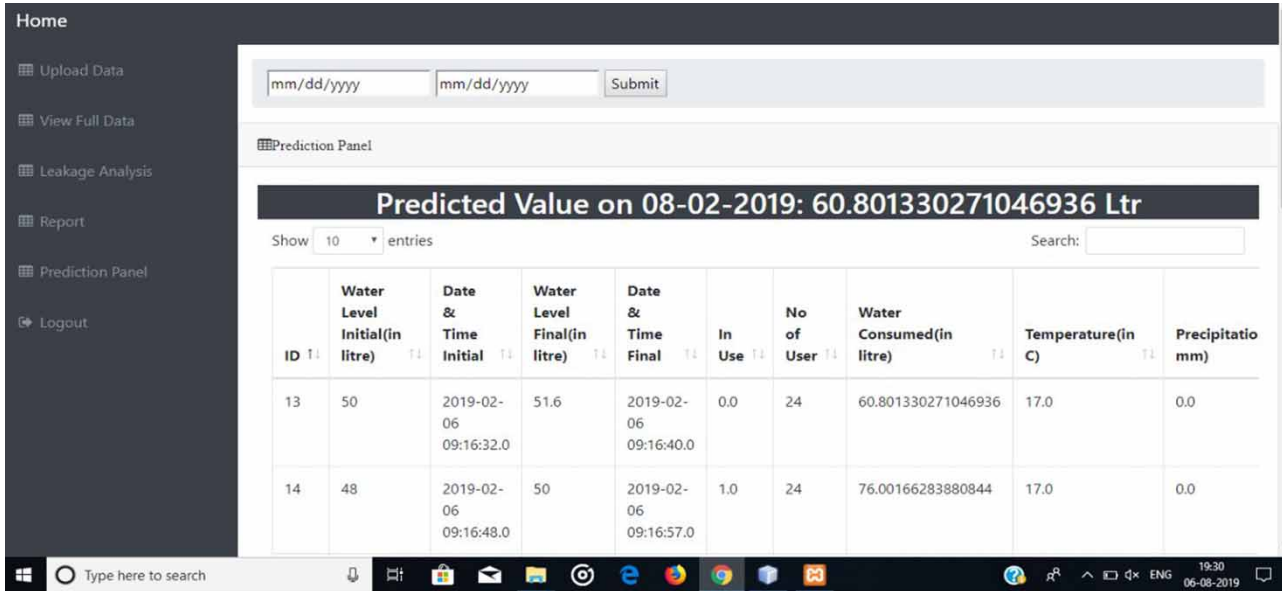


Figure 5 | Prediction of water for next day.

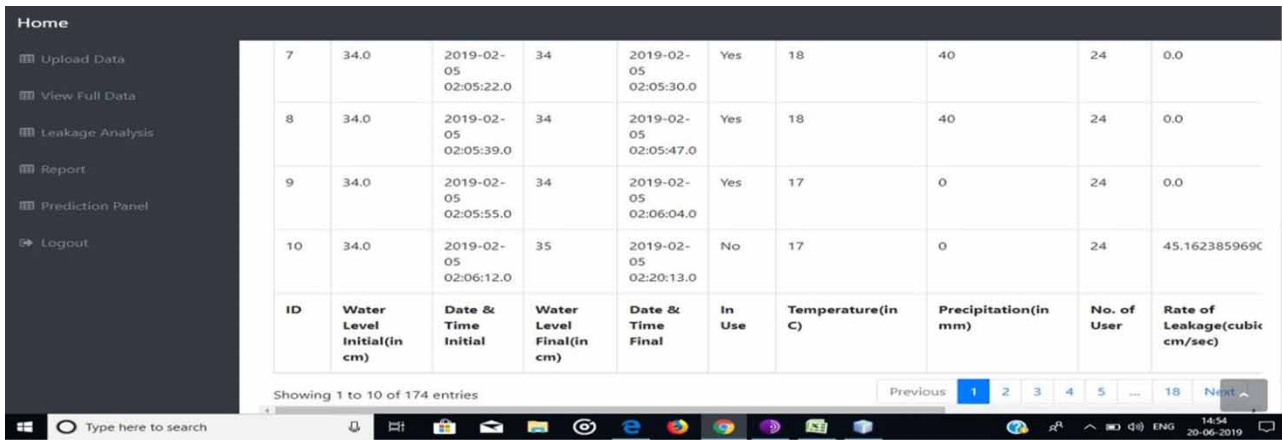


Figure 6 | Leakage analysis interface 2.

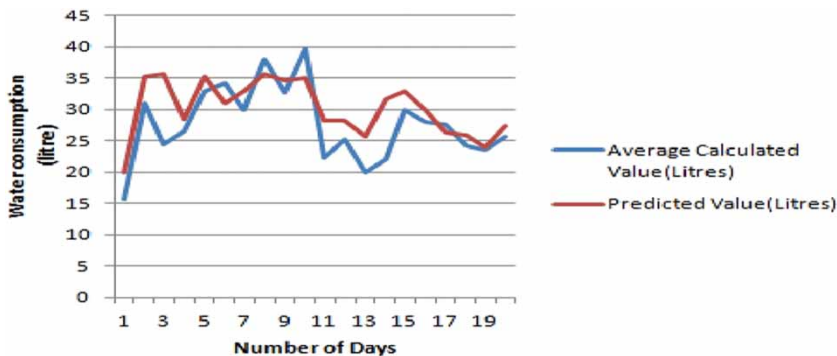


Figure 7 | Graph between average calculated values and predicted values.



Figure 8 | Graph showing error on various dates.

been detected or not and, if detected, then the rate of leakage is displayed.

RESULTS AND DISCUSSION

Figure 7 shows both the average calculated values and the predicted values of water consumption versus the number of days. The result is susceptible to many factors, such as holidays, temperature, weather, etc. Figure 7 clearly shows that the predicted values and the calculated average values tend to coincide as the data value increases.

The difference between the predicted and the actual values, i.e. the error value graph, is as shown in the Figure 8. We found a maximum error of 46.75%, which was reported at the beginning when the training data set was small. The average error is around -11.34% . The minimum value of error is as low as -1.856% . As the training data set increases, the error value approaches to minimum.

CONCLUSIONS

Water wastage has become a huge problem and therefore this paper proposes an IoT-based water management system for collecting data, predicting average requirement of water per person in a household and in society, and monitoring their consumption. An algorithm to detect leakage has been proposed and can be used to calculate the rate of leakage. A web interface allows the user to visualize the water usage. The results show that the error between the predicted and actual values is reduced as the size of data set increases.

Future scope

If data are collected continuously over a year, then average daily, weekly, monthly, and yearly consumption of water

can be calculated and more accurate water consumption values can be predicted.

If large data set values are recorded, it will be possible to identify a connection between water consumption and events such as weather, holidays, and so on.

Application

The system is highly applicable to urban water management. It can be used to determine water requirements in a society or complex; to calculate the supply and demand relationship; and to detect leakage and leakage rate, allowing immediate action to be taken to correct the cause of leakage, and ultimately saving water.

Advantages of the system

1. It saves water and reduces its loss due to unavailability of proper checks.
2. It saves a lot of users' efforts when checking the level in the water tanks.
3. It allows an equal distribution of water for all (Chatterji & Haq 2019).

ACKNOWLEDGEMENTS

This research work has been supported by JSSATE, NOIDA. We also thank its staff members for helping us collect the water data.

REFERENCES

- Adedeji, K., Hamam, Y., Abe, B. & Abu-Mahfouz, A. 2017 *Leakage detection and estimation algorithm for loss reduction in water piping networks*. *Water* 9 (10), 773.
- Alvisi, S., Luciani, C. & Franchini, M. 2019 *Using water consumption smart metering for water loss assessment in a DMA: a case study*. *Urban Water Journal* 16, 1–7.
- Boyle, T., Giurco, D., Mukheibir, P., Liu, A., Moy, C., White, S. & Stewart, R. 2013 *Intelligent metering for urban water: a review*. *Water* 5 (3), 1052–1081.
- Britton, T. C., Stewart, R. A. & O'Halloran, K. R. 2013 *Smart metering: enabler for rapid and effective post meter leakage identification and water loss management*. *Journal of Cleaner Production* 54, 166–176.
- Bureau of Waterworks, Tokyo Metropolitan Government 2018. *Prevention of leakage in Tokyo*. Available from:

- <http://www.waterprofessionals.metro.tokyo.jp> (accessed 9 November 2018).
- Candelieri, A., Soldi, D. & Archetti, F. 2015 Short-term forecasting of hourly water consumption by using automatic metering readers data. *Procedia Engineering* **119**, 844–853.
- Chatterji, S. & Haq, Z. 2019 In PM Modi's water guarantee plan, expect 55 litres/day per person. Available from: <https://www.hindustantimes.com/india-news/in-water-guarantee-plan-55-litres-day-per-person/story-cBlvjkOYY664ZurR9EjYTM.html> (accessed 30 August 2009).
- Cominola, A., Giuliani, M., Piga, D., Castelletti, A. & Rizzoli, A. E. 2015 Benefits and challenges of using smart meters for advancing residential water demand modeling and management: a review. *Environmental Modelling & Software* **72**, 198–214.
- El-Zahab, S. & Zayed, T. 2019 Leak detection in water distribution networks: an introductory overview. *Smart Water* **4** (1), 5.
- Gama-Moreno, L. A., Corralejo, A., Ramirez-Molina, A., Torres-Rangel, J. A., Martinez-Hernandez, C. & Juarez, M. A. 2016 A design of a water tanks monitoring system based on mobile devices. In: *2016 International Conference on Mechatronics, Electronics and Automotive Engineering (ICMEAE)*. IEEE, pp. 133–138.
- Gautam, J., Patil, G. M., Gupta, A., Agarwal, A., Shah, S. & Mittal, S. 2018 IOT based real time monitoring of water levels in tanks using machine learning and android app. In: *Proceedings IWA Water Reuse 2018*, October 2018, Phuket, Thailand.
- Ghalekhondabi, I., Ardjmand, E., Young, W. A. & Weckman, G. R. 2017 Water demand forecasting: review of soft computing methods. *Environmental Monitoring and Assessment* **189** (7), 313.
- Gurung, T. R., Stewart, R. A., Sharma, A. K. & Beal, C. D. 2014 Smart meters for enhanced water supply network modelling and infrastructure planning. *Resources, Conservation and Recycling* **90**, 34–50.
- Ishido, Y. & Takahashi, S. 2014 A new indicator for real-time leak detection in water distribution networks: design and simulation validation. *Procedia Engineering* **89**, 411–417.
- Jiang, B., Mu, T., Zhao, M., Shen, D. & Wang, L. 2017 Study of PSO-SVM model for daily water demand prediction. In: *2017 5th International Conference on Mechatronics, Materials, Chemistry and Computer Engineering (ICMMCCE 2017)*. Atlantis Press, Amsterdam, The Netherlands.
- Luciani, C., Casellato, F., Alvisi, S. & Franchini, M. 2019 Green smart technology for water (GST4Water): water loss identification at user level by using smart metering systems. *Water* **11** (3), 405.
- Maevsky, D., Bojko, A., Maevskaya, E., Vinakov, O. & Shapa, L. 2017 Internet of things: hierarchy of smart systems. In: *2017 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, Vol. 2. IEEE, pp. 821–827.
- Mutchek, M. & Williams, E. 2014 Moving towards sustainable and resilient smart water grids. *Challenges* **5** (1), 123–137.
- Ng, A. 2020 Machine Learning. Available from: <http://openclassroom.stanford.edu/MainFolder/DocumentPage.php?course=MachineLearning&doc=exercises/ex8/ex8.html> (accessed 24 January 2020).
- Peña-Guzmán, C., Melgarejo, J. & Prats, D. 2016 Forecasting water demand in residential, commercial, and industrial zones in Bogotá, Colombia, using least-squares support vector machines. *Mathematical Problems in Engineering*, **2016**, 1–10.
- Ray, S. 2017 Understanding Support Vector Machines Algorithm from examples (along with code). Available from: <https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/> (accessed 9 July 2018).
- Sayre, S. S. & Taraz, V. 2019 Groundwater depletion in India: social losses from costly well deepening. *Journal of Environmental Economics and Management* **93**, 85–100.
- Shahanas, K. M. & Sivakumar, P. B. 2016 Framework for a smart water management system in the context of smart city initiatives in India. *Procedia Computer Science* **92**, 142–147.
- Vijai, P. & Sivakumar, P. B. 2016 Design of IoT systems and analytics in the context of smart city initiatives in India. *Procedia Computer Science* **92**, 583–588.
- Vijayakumar, N. & Ramya, A. R. 2015 The real time monitoring of water quality in IoT environment. In: *2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*. IEEE, pp. 1–5.
- Willis, R. M., Stewart, R. A., Panuwatwanich, K., Williams, P. R. & Hollingsworth, A. L. 2011 Quantifying the influence of environmental and water conservation attitudes on household end use water consumption. *Journal of Environmental Management* **92** (8), 1996–2009.
- Xu, Y. & Goodacre, R. 2018 On splitting training and validation set: a comparative study of cross-validation, bootstrap and systematic sampling for estimating the generalization performance of supervised learning. *Journal of Analysis and Testing* **2** (3), 249–262.

First received 5 November 2019; accepted in revised form 24 February 2020. Available online 10 March 2020