# Urban Water Quality Prediction


smartbridgeinternship

Machine Learning for Engineer

Mr.Surya Tej Mahankali/Mr.Rammohan Bethi

M. Manyu

V.Abhinaya

A.Kavya

I.Prudhvi raj


## 1.Introduction

### 1.1 Overview

With the rapid development of economy and accelerated urbanization, water pollution has become more and more serious. Urban water quality is of great importance to our daily lives. Prediction of urban water quality help control water pollution and protect human health. To overcome this kind of problem statement, we developed a deep learning model to predict the water quality.Therefore, understanding the problems and trends of water pollution is of great significance for the prevention and control of water pollution. We have proposed a system that uses Machine learning algorithms to predict the water quality in Urban & to forecast the predictions.

### 1.2Purpose

Prediction of urban water quality help control water pollution and protect human health. However, predicting the urban water quality is a challenging task since the water quality varies in urbanspaces non-linearly and depends on multiple factors, such as meteorology, water usage patterns, and land uses. In this work, we forecast the water quality of a station over the next few hours from a data-driven perspective, using the water quality data and water hydraulic data reported by existing monitor stations and a variety of data sources we observed in the city, such as meteorology, pipe networks, structure of road networks, and point of interests (POIs). We evaluate our method with real-world data sets, and the extensive experiments verify the advantages of our method over other baselines and demonstrate the effectiveness of our approach.

## 2.Literature survey

### 2.1 Existing Problem

Increasing population and climatic variation driven by climate change has led to water scarcity across world. As cited in United Nations Environmental program me 2002, by 2025, 1.8 billion people will be living in countries or regions with absolute water scarcity about two-third of the world population, mainly in developing countries will face moderately to high water stress and half of the population will face problem due to water scarcity. Ongoing mismanagement of resources and wasteful behavior in India has led to the over exploitation of water resources, particularly groundwater. A large part of India fall under the category of physical water scarcity where availability of natural water resources is not enough to ensure the future water needs hence they need to increase their efficiency of water use and wisely maintain their available water resources. The concentration of water related problems will be more in urban settlements where quantity of available water reducing day by day. The potable water as a commodity is highly delicate and vulnerable to pollution and contamination. As such it has to be handled with a high degree of care. This paper focuses on issues and challenges associated with urban water requirements and suggest pragmatic solutions and policy measures to ensure clean, continual and convenient access to drinking water supply to all the residents of urban settlements.

Most of the world countries are facing the problem of fresh water scarcity mainly due to increasing population and climatic variation in rainfall is driven by climate change (Image 1). About two-third of world countries mainly developing countries will face moderate to high water stress water and half of the total world population will face real water constraints by 2025 (United nations environmental program me 2002). Many of the European countries in the temperate zone having plentiful of fresh water resources are also facing the shortage of water supply due to successive water droughts driven by climate variations lead to drying of many water resources and water level in aquifers have reached to the critical point (V. Lazarus et. al. 2001). Large part of India also fall under the category of physical water scarcity where availability of natural water resources is not enough to secure their future water needs hence they need to increase their efficiency of water use and wisely maintain their available water resources (Seckler et al. 2008; Christine L. Moe et. al. 2006).

In India, right to fresh water for personal and domestic uses is not mentioned explicitly in Indian constitution but clean and affordable water is essential to life and one of the fundamental human rights protected under international human rights law. Freshwater is a finite resource and is also a basic requirement for human body. Water is used mainly for domestic, agricultural and industrial purposes and also food production are essentially a function of water level at farm and industrial levels (FAO 2008). The greatest demand of freshwater resources is in agriculture for food production as about 70% of the developed water supplies used in irrigation (Seckler et el, 1998; Christine L. Moe et al , 2006). About 300 to 3000 litre water required to produce 1kg of food grain and that food production for a balanced diet requires 1300cubic of water per person per day.

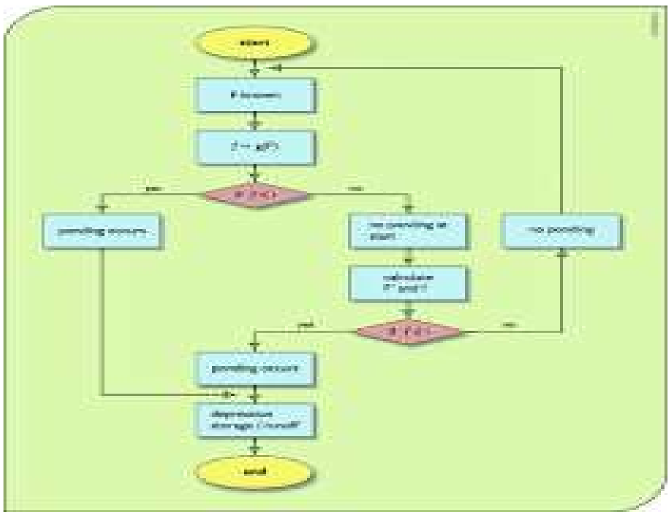
## 2.2 Proposed problem

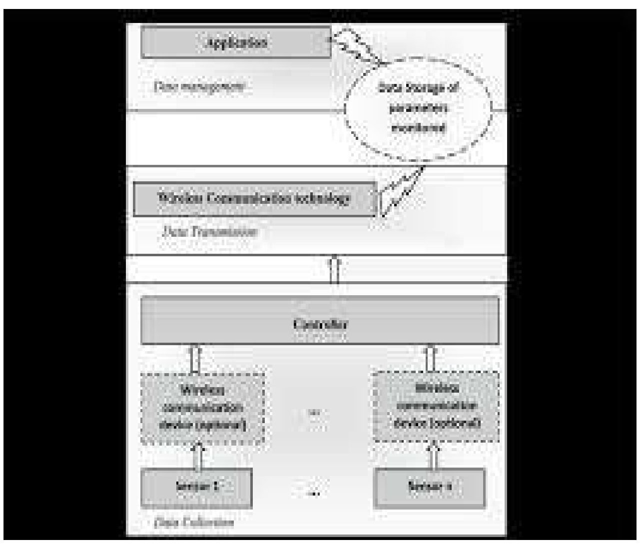
Like water quantity, water quality also has been a matter of political agenda across the world and India, Delhi is the best example of the same. In the absence of strong regulation, industrial and domestic waste water is discharged in rivers, canals and underground water sources. About 70% of underground and surface water resources in India have been contaminated (Sudhakar M.Rao et al, 2004). High levels of pollution of the groundwater has been caused by the printing and dyeing units in Pali and Jodhpur in Rajasthan, Jetpur and Rajkot in Gujarat, Tannery industry in North Arcot district in Tamil Nadu. Printing and dyeing units in Panipat and Sonipat in Haryana are among other examples.

Therefore, understanding the problems and trends of water pollution is of great significance for the prevention and control of water pollution. We have proposed a system that uses Machine learning algorithms to predict the water quality in Urban & to forecast the predictions.

## 3.Theoritical Analysis

### 3.1Block diagram

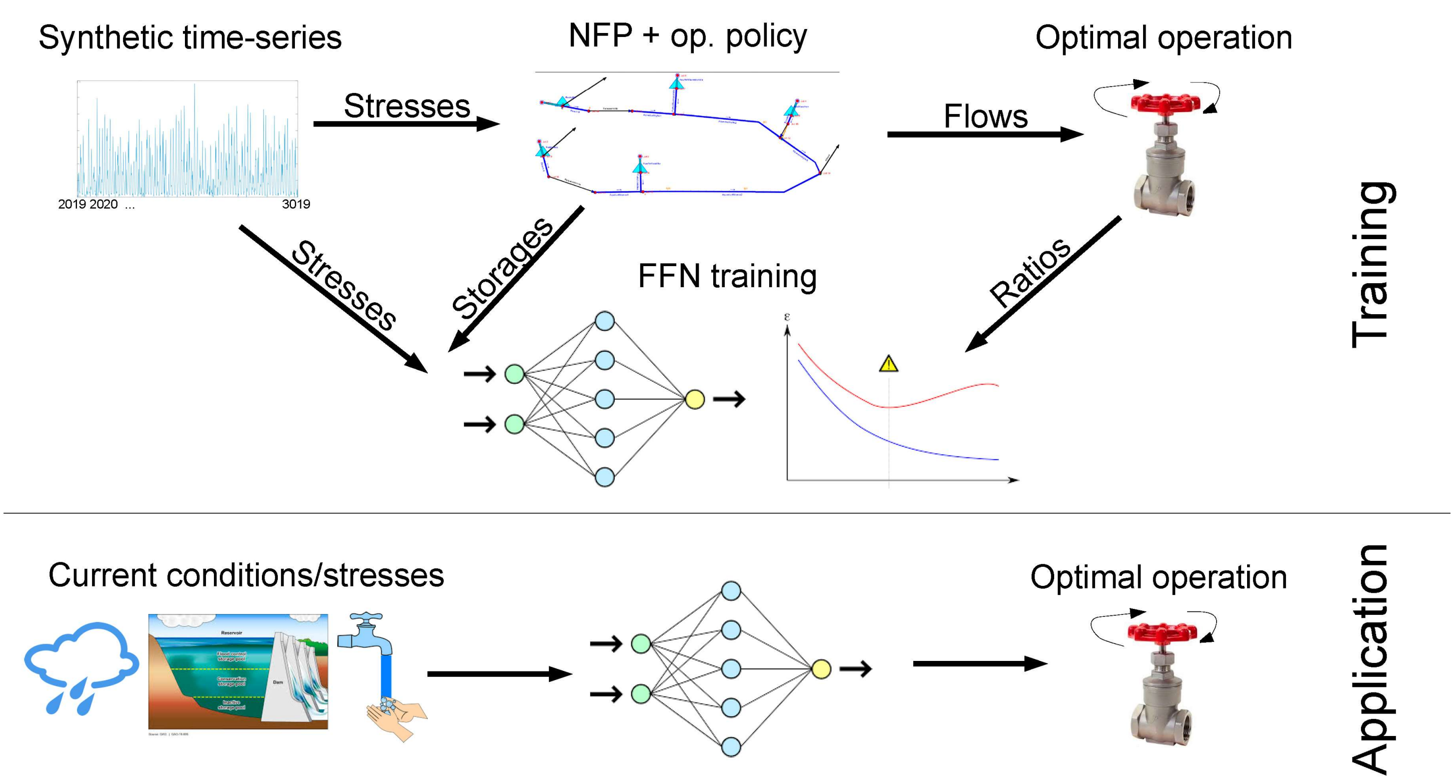





### 3.2 Hardware and software designing

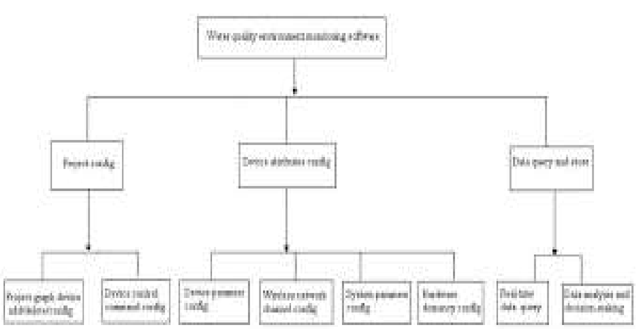
The functions of water quality monitoring software in application layer are mainly the water quality data analysis and decision-making. The software development tool is Visual studio 2005, and develops, compile, debug in Windows XP. The software architecture based on .NET framework, using MVC structure, implements the separation of logic layer, display layer and data layer, which improved application scalability and reusability.In this architecture, the software hierarchical structure has three-layers: display layer, logic layer and data layer. Data acquisition and device control layer is related to the wireless sensor network hardware of the system architecture. The whole workflow process of the software is: firstly each sensor node in the wireless sensor network sends the water quality monitoring data to the water quality monitoring software; secondly in the design project, the users can see the real-time data by double-click the corresponding graphical device, and get the network topology. The user also can get the trends of one sensor data in a period or compare several sensors data in one comparison chart; thirdly according to the water quality data and device status, the users can add/delete device information, change the parameters of the device channel and set the sleeping status to the device for saving the battery power through the device configure module; fourthly through data analysis module, the software calls the knowledge from the knowledge database in the data layer to analyze the water quality data and decides to whether send control commands to the associated devices. The users can real-time monitor and manage aquatic water quality conveniently and efficiently by the various monitoring and control functions that the entire software 6 modules work together to achieve.

The software main function modules include project design module, device configure module and data




query/store/analyze module. Project design module through the user interface can display real-time data of water quality parameters in the monitoring region, over-limit alarm and historical data records; Device configure module uses communication protocol to communicate with the routing node, in order to achieve data transmission and device configuration aims, and in this module sensor network topology can be displayed to users; Data module uses database to implement the storage of the history monitoring data, and provides query and analysis operations to the users. Water quality monitoring software functions structure.

As a water quality monitoring software, the configuration of the used devices would help users design and management of the project, so the module uses configuration model on the graphical devices in the development. The graphical devices configuration in the water quality monitoring software include ponds, access node, routing node, collection node, control valves, switches. Users can design the project by dragging and dropping the graphical device from the graphical devices configuration toolbar. All graphical devices in the project can be moved, added relations, and viewed the real-time data by clicking in the main project interface. For the collection node and routing node, because they communicate with the software by



protocol, the communication state maybe not stabilization in the project. The corresponding graphical device has two states, one green color expresses the good communication state, and the other red color expresses the bad state. Because the communication state refresh in time, the users can manage the device communication state from the project main interface. Also the control device has two state, green expresses the valve open, red expresses close. The control device can be set auto control to implement the operations depend on the monitoring data of the sensor. Users can re-design new projects, and user-designed project can be saved to reuse that implement different management in different projects for the user.

## 4.Experimental Investigation

The sample water was collected into precleaned, airtight plastic pots fro the lake in the evening during February to March 2017 and transported to the environmental engineering laboratory in theInternational Islamic University Malaysia for inorganic solids experi-mentation of the water sample. A s a result of its geographical lo-cation, Malaysia does not exhibit any significant weather changes throughout the year. Thus, it was deemed that the February to March period may represent the whole year. The water samples were collected at four different times during February–March and their physicochemical characteristics analysed. The water samples were obtained at two-week inter vals, for a total of four sampling efforts, two being during the dr y time and two after precipitation events. Theaverage result of the four water analyses for each pa-rameter was calculated. The investigated parameters included pH, dissolved oxygen (DO), chemical oxygen demand (COD), biological oxygen demand in 5 days (BOD5) and turbidity, inorganic solids, total dissolved solids (TDS), total suspended solids (TSS), volatile suspended solids ( VSS), copper(Cu), zinc(Zn), total phosphorus (TP) and total nitrogen (TN). Turbidity and total suspended solids were the only physical characteristic of the water examined.

**P$^{H}$ and dissolved oxygen (DO) concentration**

Twenty-five millilitres of water samples was obtained from Taman Idaman L ake. The pH and DO were measured in situ with a YSI meter. The pH electrode was washed with distilled water. The electrode was submerged into the 25 ml water sample and stirred thoroughly. The reading was taken only after the measured pH value was stable with the same range. The meter also measures the dissolved oxygen concentration on the basis of the rate of diffusion of molecular oxy-gen across a membrane, with the same steps being used to measure the DO as for the pH.

**COD and BOD5**

The chemical oxygen demand (COD) was assayed by the closed reflux method, while the BOD5 was analysed by the five-day test method. The COD test incorporated the complete degradation of biodegradable organics via utilization of strong oxidizing agents (i.e., potassium dichromate). A 100 ml water sample was homogenized. A COD reactor was preheated to 150°C and a plastic shield placed in front of the reactor. A 0.2 ml reagent was pipetted into the sample,
heated for 2 hr and cooled in a water bath. The tube was then ana-lysed with a spectrophotometer. Volatile organic compounds (VOC) were completely oxidized in this system, compared to the open sys-tem, because of a longer contact time with oxidants (HATCH, 2017).Dilution for BOD5 measurement is necessary when the quantity of DO consumed by microorganisms is greater than the available DO in the air-saturated BOD5. The needed sample size is determined
from the anticipated BOD5 value, which is generally half of the COD value.

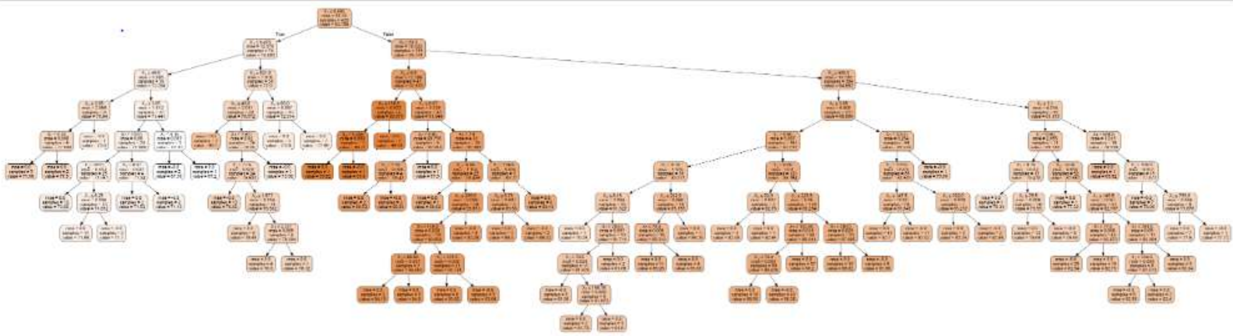
The sample dilution water was prepared by measuring 50 ml of sample water into BOD bottle containing nutrient buffer. The bottle was sealed with sealing gel to prevent gaseous interferences and placed in an incubator at 20 ± °C. The initial DO content was recorded. The DO content for every day thereaf ter was recorded for 5 consecutive days, using an automated DO probe.

## Copper (Cu), zinc (Zn), total phosphorus (TP) and total nitrogen (TN)

For copper determination, sachets containing reagent ZincoVer 5 was emptied into a vial containing 10 ml sample water. After dissolution of the powder, 0.5 ml of cyclohexanone was added to the mixture and shaken vigorously for 30 s. The sample was then allowed to react for 3 min, after which it is measured with a spectrophotometer. A blank sample was used to compare the spectrophotometer readings.For zinc determination, a sachet containing CuVer 1 copper reagent powder pillow was emptied into a vial containing 10 ml of water sample. After a two-minute reaction time, the sample was measured with a spec trophotometer. A blank sample was used to compare the spectrophotometer readings.To determine the total nitrogen (TN), a COD reactor was set to 103–106°C, and one persulphate reagent powder pillow was added to a total nitrogen hydroxide reagent vial, which was placed in the COD reactor for 30 min. After cooling, a total nitrogen reagent A powder pillow was added to the vial and shaken in order to start the reaction. A TN reagent B powder pillow was then added to the vial, which was shaken again to facilitate the reaction. Two millilitres of the digested sample was shifted to a vial of (TN) reagent C , and an-other two millilitres of treated digested blank was added to another vial of the (TN) reagent C. The vials were inverted several times to mix and lef t for 5 min to react. Both vials were wiped off and measured in a spectrophotometer, with the TN readings being recorded.

| | Unnamed: 0 | station | do | ph | co | bod | na | tc | wqi |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 1393.0 | 6.7 | 7.5 | 203.0 | 1.8965 | 0.100 | 27.0 | 93.82 |
| 1 | 1 | 1399.0 | 5.7 | 7.2 | 189.0 | 2.0000 | 0.200 | 8391.0 | 76.96 |
| 2 | 2 | 1475.0 | 6.3 | 6.9 | 179.0 | 1.7000 | 0.100 | 5330.0 | 79.28 |
| 3 | 3 | 3181.0 | 5.8 | 6.9 | 64.0 | 3.8000 | 0.500 | 8443.0 | 69.34 |
| 4 | 4 | 3182.0 | 5.8 | 7.3 | 83.0 | 1.9000 | 0.400 | 5500.0 | 77.14 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 1986 | 1986 | 1330.0 | 7.9 | 738.0 | 7.2 | 2.7000 | 0.518 | 202.0 | 72.06 |
| 1987 | 1987 | 1450.0 | 7.5 | 585.0 | 6.3 | 2.6000 | 0.155 | 315.0 | 72.06 |
| 1988 | 1988 | 1403.0 | 7.6 | 98.0 | 6.2 | 1.2000 | 0.516 | 570.0 | 66.44 |
| 1989 | 1989 | 1404.0 | 7.7 | 91.0 | 6.5 | 1.3000 | 0.516 | 562.0 | 66.44 |
| 1990 | 1990 | 1726.0 | 7.6 | 110.0 | 5.7 | 1.1000 | 0.516 | 546.0 | 66.44 |

## 5.Flowchart:



## 6.Results

### Linear regression:

```
from sklearn.metrics import r2_score
r2_score(y_test,y_pred)
```

```
0.6120501654221961
```

```
mr.predict([[1393.0,6.7,7.50,203.0,1.8965,0.100,27.0]])
```

```
array([89.64148702])
```

### Multi-linear regression:

```
from sklearn.metrics import r2_score
r2_score(y_test,y_pred)
```

```
0.6120501654221961
```

```
mr.predict([[1393.0,6.7,7.50,203.0,1.8965,0.100,27.0]])
```

```
array([89.64148702])
```

### Decision Tree regression:

```
r2_score(y_test,y_pred1)
```

```
0.9387170993756437
```

```
dataa.predict([[1393.0,6.7,7.50,203.0,1.8965,0.10
```

```
array([93.82])
```

### Random forest regression:

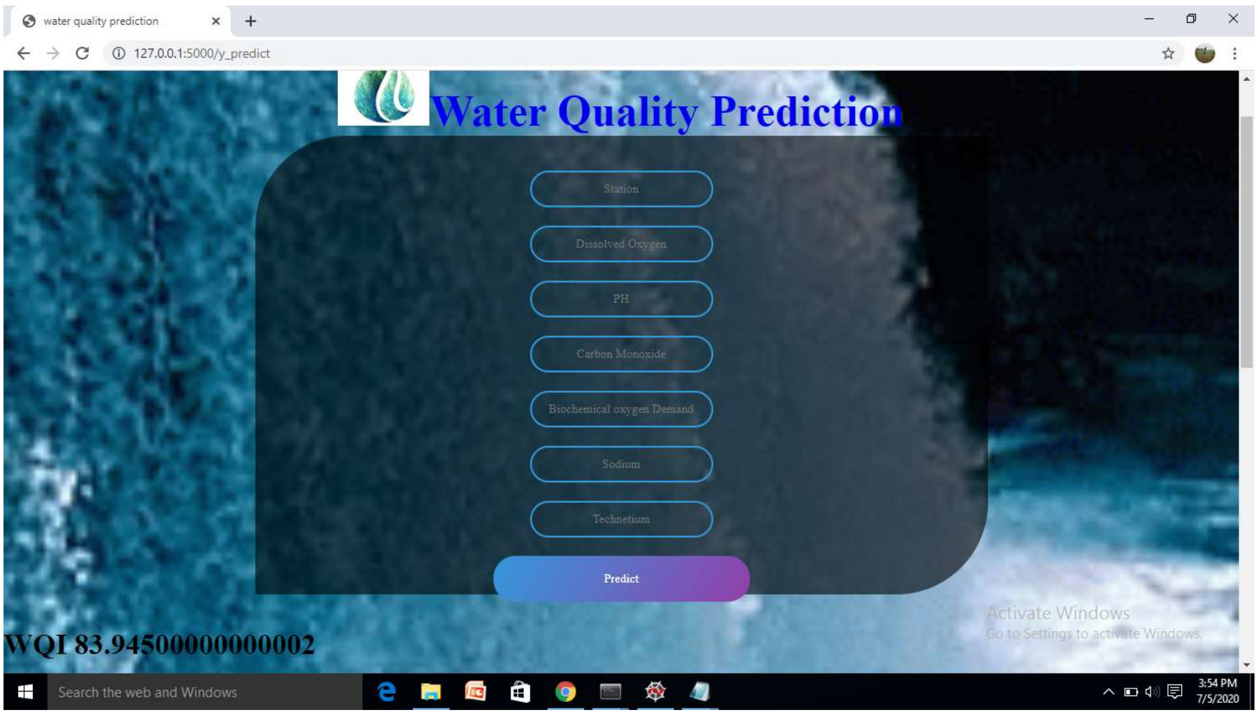
```
#evaluation
from sklearn.metrics import r2_score
r2_score(y_test,y_predict)
```

```
0.9419559375806802
```

```
rf.predict([[1393.0,6.7,7.50,203.0,1.8965,0.100,27.0]])
```

```
array([92.995])
```

## Flask UI:




## 7.Advantages and Disadvantages

### pros:

- Leads to Better Health. **Quality water** helps prevent waterborne illness. ...
- Saves Money on Plumbing Fixtures and Appliances. ...
- Fresher Taste and Odor. ...
- Less Cleaning. ...
- Maintain Your **Water** Softener. ...
- Keep Your **Water** Filters Clean. ...
- Wash Your **Water** Containers.

### cons:

- Most cover crops don't generate income.Can be tougher to grow

in cold climates if not planted immediately after harvest.Some farmers fear a short spring will mean not enough time to kill cover before planting corn.

- Takes land put of production.Doesn't filter water from underground drainage tiles.
- Relatively expensive at outset.During heavy rain,not all water is filtered.
- Farmers may need different

equipment,more herbicide with no-till.

- Relatively expensive to construct.While terraces stop soil movements,they don't always decrease water pollution.
- During heavy rains,some water will go around bioreactors and not be filtered.Doesn't improve soil.

## 8.Applications

### Agriculture

Clearly, water quality is critical to crop production. Poor quality water in general can slow plant growth and negatively impact appearance. In addition, certain plants grow better when certain characteristics are present in irrigation water; the same is true for the converse—plant growth may be hindered by the presence of certain factors. For example, water with a too-high pH level (an indicator of acidity) could be detrimental to some plants, making it difficult for them to absorb nutrients from the soil; others are less affected by soil pH and are able to absorb higher levels of aluminum or magnesium in even very acidic water.

Another issue relates to salt tolerance. The nutrients plants use for development are absorbed as ionic salts; however, some plants tolerate high salinity while others will not. Irrigation water ranges in salinity; that salinity is usually measured as the electrical conductivity of water because dissolved salts in water conduct electricity. The higher the conductivity, the higher the salt concentration. Depending on the type of crop you're trying to grow, a smart water quality monitoring system can help maintain the optimum conductivity level for maximum growth.

There's significant economic value in monitoring water quality for the ag space. It can be an extremely useful tool for managing irrigation water, and help farmers maximize crop production.

### 2. Drinking Water

Many cities throughout the U.S. use water quality monitoring equipment continuously to evaluate the safety of their water supplies. (The recent water woes of Flint, Michigan, should serve as an important reminder to municipal governments that water quality cannot be overlooked.)

Drinking water is typically tested for the presence of bacteria, so it's critical to examine the sterility of the water—can it support life? The oxidation-reduction potential (ORP) essentially measures the ability of water to break down contaminants; therefore, some municipalities use ORP sensors to measure the dissolved oxygen in water. High ORP levels indicate that the water is not capable of supporting bacteria, and so are useful in determining the overall quality of a particular water supply source.

### 3. Aquaculture

Aquaculture, or the cultivation of aquatic plants and fish, requires high-quality water to promote production and increase profitability. Many facilities now use water quality monitoring equipment to measure water temperature and conductivity, two of the many factors that can affect the physical condition of aquatic animals.

Optimal water quality varies by species, so the parameters will vary across the industry. Water quality parameters commonly monitored among farming systems include temperature, dissolved oxygen, pH, alkalinity, hardness, ammonia, and nitrites. And depending on the culture system, carbon dioxide, chlorides, and salinity may also be monitored.

### 4. Water Loss Detection

Just about every building has an invisible network of pipes that carries water throughout the facility. For the most part, it's difficult to detect leaks in these pipes, and even when they're found, it's usually after much water has already been lost. Pipe leakage is responsible for a significant amount of water loss in both industrial/commercial facilities and public water systems, too. That's important not just from a cost perspective, but also from a sustainability perspective—water conservation is both good for the environment and good for business.

Water flow sensors measure the flow of water through a pipe and its rate of change. Pulse meters are helpful in understanding the rate of normal flow, and therefore present a way to detect when the normal rate changes. In manufacturing facilities, flow sensors can be used to measure the volumetric flow rate of any liquid, gas, or steam. Changes in flow rate may be an indicator of pipe leakages or other operational malfunctions, giving building managers a chance to address problems before too much water (or any other liquid resource) is wasted.

## 9.Conclusion

Water quality monitoring is of vital imperativeness as it gives particular data about the nature of water. Different sorts of physical, substance and organic data are joined to determine at a differing qualities list. The composite impact of huge physical, synthetic and natural parameters is reflected in the index. Diversity indices for water quality, monitoring can be used for resources allocation, location ranking, standard enforcement, trend analysis, public information and scientific research. The present work describes the dominant physico–chemical variables and their relation with phytoplankton density and community along a pollution gradient. Study was also done on the surfaceplankton populace in the oceanic environment of Waghur and Tapi waterway water. The mechanical effluents from different commercial ventures in and around Jalgaon contain various poisonous substances once went into the river Waghur and Tapiaffecting the water quality. As an outcome, the planktonpopulation of the Waghur and Tapi River has been influenced in wording ofabundance and differing qualities by evaluatingthe microscopic fish file as the water quality criteria with reference tofreshwater bodies dirtied by different industrial and domestic actions. Species diversity is of considerable importance in assessing the extent of damage to natural system by human interference severe disturbances cause a marked decline in the diversity. Biodiversity, a concept used to describe the dimension of living organisms in a given area, takes into consideration mixture of the life structures, the qualities they contain and the environment they are found. However, there

are no registered effluents that are forced into the water bodies at the stretches of rivers in North Maharashtra Region. But small open drainages from the bank of river are big in number that they have to consider as non-point sources of pollution. Indiscriminate defecation, littering of solid wastes all around and into the rivers, cremation and dumping of carcasses, mass bathing, cattle wallowing and agriculture washouts of pesticides, fertilizers and 215 insecticides are a bit anthropogenic stresses which have been studied in detailed of river Waghur and Tapi. Narrow fluctuations in air temperature across seasons revealed that the dominant climatic factor in this temperate high altitude river system is not air temperature. A correlation of precipitation between distinctive seasons unmistakably demonstrated that precipitation is the major climatic component commanding in the framework.

## 10.Future Scope

Based on the finding of the present investigation the following suggestions may be made for future studies:

1. All biological parameters of water can be studied to know the quality of water.
2. The study of air pollution can be studied. After that the effect of air pollution on quality of water can be studied.
3. Development and evaluation of a portable filter can be undertaken for its implementation in rural areas.
4. The ANN method is excellent in prediction of future data basing on the previous data provided the number of datas are more. So suitable methods can be developed to predict any physico-chemical parameter for any future years.
5. Future studies including a research on the effect of temperature on the chemisorptions mechanism and the behaviour of the culture to regenerate the resin are recommended.
6. It is also suggested to study the effect of decreasing the ratio of resin mass to volume of regenerant solution since it could provide valuable

information for optimizing the process.
7. Perchlorate reduction in groundwater by Thiobacillus ferroxidant bacteria should be done.


## Bibilography

1. Ahern, J. (2007) Green infrastructure for cities: The spatial dimension, a chapter in Cities of the Future: Towards integrated sustainable water and landscape management.
1.1 Brown , IWA Publishing, London, UK,Barnard, J.L. (2007) Elimination of eutrophication through the resource recovery ,The 2007 Clarke Lecture, National Water Research Institute, Fountain Valley, CA.
1.2 Brown, P. (2007) The importance of water infrastructure and environment in tomorrow's cities, in Cities of the Future: Towards Integrated Sustainable Water and Landscape Management (V. Novotny) , IWA Publishing, London, UK.
2.Brundtland, G. (ed.) (1987) Our Common Future: The World Commission on Environment and Development, Oxford University Press, Oxford.
3.Elkington, J. (1997) Cannibals with Forks: The Triple Bottom Line of 21st century Business, Capstone Publishing, Oxford
4.Engle, D. (2007) Green from top to bottom, Water Efficiency 2(2):10-15 Forman, R.T,.T. et al. (2003) Road Ecology: Science and Solutions, Island Press, Washington.
5.Furumai,H.Reclaimed stormwater and wastewater and factors affecting their reuse, in Cities of the Future: Towards integrated sustainable water and landscape management (V. Novotny and P. Brown, eds.), IWA Publishing, London, UK.
6.IPCC (2007). Summary for Policy Makers, Climate Change 2007: The Physical Scientific Basis, Fourth Assessment Report, Intergovernmental Panel on Climatic Change, Working Group (WG 1),Geneva.
7.International Organization for Standardization (ISO) (2006) Environmental Management – Life – Cycle Assessment – Principles and

Framework, International Standard 14020, Geneva,Switzerland.
8.Heaney,Centralized and decentralized urban water, wastewater & stormwater systems,in Cities of the Future:Towards integrated sustainable water and landscape management (V. Novotny and P. Brown, eds.), IWA Publishing, London, UK.
9.Hammer, T.R. (1972) Stream channel enlargement due to urbanization, Wat. Res. Research 8:139-167.
10.Hill,K.(2007)Urban ecological design and urban ecology: an assessment of the state of current.knowledge and a suggested research agenda,in Cities of the Future: Towards integrated sustainable water and landscape management (V. Novotny and P. Brown, eds), IWA Publishing, London, UK.